U.S. DEPARTMENT OF COMMERCE
Patent and Trademark Office

SEARCH REQUEST FORM

Requestor's Name: _____ Serial Number: _____

Date: _____ Phone: _____ Art Unit: _____

Search Topic:
Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: 05-28-03
Searcher: Rowley E 9-9-95
Terminal time: 20
Elapsed time: _____
CPU time: _____
Total time: 25
Number of Searches: _____
Number of Databases: 1

Search Site
_____ STIC
_____ CM-1
_____ Pre-S

Type of Search
_____ N.A. Sequence
_____ A.A. Sequence
_____ Structure
_____ Bibliographic

Vendors
_____ IG Suite
_____ STN
_____ Dialog
_____ APS
_____ Geninfo
_____ SDC
_____ DARC/Questel
_____ Other (Specify)

Jiang, D.
09/930,312
Seq. IDs 1 & 2

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 18:02:22 ; Search time 1902 Seconds
(without alignments)
4911.672 Million cell updates/sec

Title: US-09-930-312-1
Perfect score: 321
Sequence: 1 ggtgtgaacttgctggcttt.....actttttaaggacaagttg 321

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		321	100.0	321	6	AX384516	AX384516 Sequence
2		321	100.0	1068	6	AX148176	AX148176 Sequence
3		321	100.0	2049	6	AX385037	AX385037 Sequence
4		321	100.0	2142	6	AX451566	AX451566 Sequence
5		321	100.0	2193	6	AX385032	AX385032 Sequence
6		321	100.0	2214	6	AX451562	AX451562 Sequence
7		321	100.0	2265	6	AX385027	AX385027 Sequence
8		321	100.0	2436	9	AF453828	AF453828 Homo sapi
9		321	100.0	2838	9	AF403384	AF403384 Homo sapi
10		233	72.6	2214	6	AX385045	AX385045 Sequence
11		233	72.6	2539	10	AF346501	AF346501 Mus muscu
12	c	220	68.5	1018	6	AX147820	AX147820 Sequence
13		220	68.5	170522	9	AL138708	AL138708 Human DNA
14		166.4	51.8	1191	6	AX128507	AX128507 Sequence
15		166.4	51.8	2274	9	AF190500	AF190500 Homo sapi
16	c	158.2	49.3	180673	10	AC068627	AC068627 Mus Muscu
17	c	158.2	49.3	213462	10	AC077689	AC077689 Mus Muscu
18		156.8	48.8	474	6	AX088165	AX088165 Sequence
19	c	148.6	46.3	164610	2	AC098990	AC098990 Rattus no
20		148.6	46.3	171869	2	AC121029	AC121029 Rattus no
21		112	34.9	187431	9	AC107219	AC107219 Homo sapi
22		109.8	34.2	58377	2	AC128964	AC128964 Rattus no
23	c	109.8	34.2	124282	2	AC121415	AC121415 Rattus no
24		109.8	34.2	173546	2	AC098607	AC098607 Rattus no
25	c	108.6	33.8	179055	2	AC122462	AC122462 Mus muscu
26		107.2	33.4	176565	2	AL136106	AL136106 Homo sapi
27	c	77.8	24.2	171869	2	AC121029	AC121029 Rattus no
28		67.2	20.9	1015	6	AX254453	AX254453 Sequence
29	c	58	18.1	173546	2	AC098607	AC098607 Rattus no
30		53.2	16.6	1080	6	AX254459	AX254459 Sequence
31		53.2	16.6	3187	6	AX254458	AX254458 Sequence
32	c	53.2	16.6	57274	2	AC017540	AC017540 Drosophil
33		53.2	16.6	163970	3	AC010580	AC010580 Drosophil
34		53.2	16.6	167396	3	AC010579	AC010579 Drosophil
35		53.2	16.6	226332	3	AE003753	AE003753 Drosophil
36		41.8	13.0	123755	2	AC120682	AC120682 Rattus no
37	c	41.8	13.0	137128	2	AC106965	AC106965 Rattus no
38		41.6	13.0	3756	3	LSGPCR	Z23104 L.stagnalis
39		41.4	12.9	154587	9	AP001208	AP001208 Homo sapi
40		41.4	12.9	163255	9	AP001330	AP001330 Homo sapi
41	c	41.4	12.9	181598	9	AC011028	AC011028 Homo sapi
42		40.8	12.7	7218	6	I66494	I66494 Sequence 14
43		40.6	12.6	163183	9	AC016574	AC016574 Homo sapi
44		40.6	12.6	203134	9	AC021077	AC021077 Homo sapi
45		40.2	12.5	321	4	AY059667	AY059667 Amblysomu

ALIGNMENTS

RESULT 1
AX384516
LOCUS AX384516 321 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 1 from Patent WO0214496.
ACCESSION AX384516
VERSION AX384516.1 GI:19577719
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Lind,P.
TITLE Novel g protein-coupled receptors
JOURNAL Patent: WO 0214496-A 1 21-FEB-2002;

```
          PHARMACIA & UPJOHN COMPANY (US)
FEATURES             Location/Qualifiers
     source          1. .321
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       76 a     67 c     59 g    119 t
ORIGIN

  Query Match             100.0%;  Score 321;  DB 6;  Length 321;
  Best Local Similarity   100.0%;  Pred. No. 6.7e-71;
  Matches  321;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60

Qy     61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120

Qy    121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180

Qy    181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240

Qy    241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300

Qy    301 AACTTTTTTAAGGACAAGTTG 321
          |||||||||||||||||||||
Db    301 AACTTTTTTAAGGACAAGTTG 321


RESULT 2
AX148176
LOCUS       AX148176                1068 bp    DNA     linear   PAT 08-JUN-2001
DEFINITION  Sequence 17 from Patent WO0136471.
ACCESSION   AX148176
VERSION     AX148176.1  GI:14347081
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1068)
  AUTHORS   Chen,R., Dang,H.T. and Lowitz,K.P.
  TITLE     Endogenous and non-endogenous versions of human g protein-coupled
            receptors
  JOURNAL   Patent: WO 0136471-A 17 25-MAY-2001;
            Arena Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1068
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      278 a    215 c    214 g    361 t
ORIGIN

  Query Match             100.0%;  Score 321;  DB 6;  Length 1068;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-71;
  Matches  321;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    589 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 648

Qy     61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 708

Qy    121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 768

Qy    181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 828

Qy    241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    829 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 888

Qy    301 AACTTTTTTAAGGACAAGTTG 321
          |||||||||||||||||||||
Db    889 AACTTTTTTAAGGACAAGTTG 909


RESULT 3
AX385037
LOCUS       AX385037                2049 bp    DNA     linear   PAT 19-MAR-2002
DEFINITION  Sequence 11 from Patent WO0214489.
ACCESSION   AX385037
VERSION     AX385037.1  GI:19578160
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Paszty,C.J., Gong,J., Daugherty,B. and Rogers,N.
  TITLE     Leucine-rich repeat-containing g-protein coupled receptor-8
            molecules and uses thereof
  JOURNAL   Patent: WO 0214489-A 11 21-FEB-2002;
            Amgen, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2049
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             1. .2049
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD28590.1"
                     /db_xref="GI:19578161"
                     /translation="MIVFLVFKHLFSLRLITMFFLLHFIVLINVKDFALTQGSMITPS
                     CQKGYFPCGNLTKCLPRAFHCDGKDDCGNGADEENCGDTSGWATIFGTVHGNANSVAL
                     TQECFLKQYPQCCDCKETELECVNGDLKSVPMISNNVTLLSLKKNKIHSLPDKVFIKY
                     TKLKKIFLQHNCIRHISRKAFFGLCNLQILILDDNPITRISQRLFTGLNSLFFLSMVN
                     NYLEALPKQMCAQMPQLNWVDLEGNRIKYLTNSTFLSCDSLTVLDLSSNTITELSPHL
                     FKDLKLLQKLDLERIEIPNINTRMFQPMKNLSHIYFKNFRYCSYAPHVRICMPLTDGI
                     SSFEDLLANNILRIFVWVIAFITCFGNLFVIGMRSFIKAENTTHAMSIKILCCADCLM
                     GVYLFFVGIFDIKYRGQYQKYALLWMESVQCRLMGFLAMLSTEVSVLLLTYLTLEKFL
                     VIVFPFSNIRPGKRQTSVILICIWMAGFLIAVIPFWNKDYFGNFYGKNGVCFPLYYDQ
                     TEDIGSKGYSLGIFLGVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVAN
                     RFFFIVFSDAICWIPVFVVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTN
                     FFKDKLKQLLHKHQRKSIFKIKKKSLSTSIVWIEDSSSLKLGVLNKITLGDSIMKPVS
                     "
     sig_peptide     1. .108
BASE COUNT      579 a    415 c    395 g    660 t
ORIGIN

  Query Match             100.0%;  Score 321;  DB 6;  Length 2049;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-71;
  Matches  321;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1570 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1629

Qy     61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1630 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1689

Qy    121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1690 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1749

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1750 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1809

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1810 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 1869

Qy     301 AACTTTTTTAAGGACAAGTTG 321
          |||||||||||||||||||||
Db    1870 AACTTTTTTAAGGACAAGTTG 1890


RESULT 4
AX451566
LOCUS       AX451566               2142 bp    DNA     linear   PAT 03-JUL-2002
DEFINITION  Sequence 5 from Patent WO0226824.
ACCESSION   AX451566
VERSION     AX451566.1  GI:21698550
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Feder,J.N., Mintier,G., Ramanathan,C.S. and Hawken,D.R.
  TITLE     A novel human g-protein coupled receptor, hgprbmy5, expressed
            highly in brain and ovarian tissues
  JOURNAL   Patent: WO 0226824-A 5 04-APR-2002;
            Bristol-Myers Squibb Company (US)
FEATURES             Location/Qualifiers
     source          1. .2142
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      613 a    435 c    408 g    686 t
ORIGIN

  Query Match           100.0%;  Score 321;  DB 6;  Length 2142;
  Best Local Similarity  100.0%;  Pred. No. 5.3e-71;
  Matches  321;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1663 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1722

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1723 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1782

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1783 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1842

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1843 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1902

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1903 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 1962

Qy     301 AACTTTTTTAAGGACAAGTTG 321
          |||||||||||||||||||||
Db    1963 AACTTTTTTAAGGACAAGTTG 1983


RESULT 5
AX385032
LOCUS       AX385032               2193 bp    DNA     linear   PAT 19-MAR-2002
DEFINITION  Sequence 6 from Patent WO0214489.
ACCESSION   AX385032
VERSION     AX385032.1  GI:19578156
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Paszty,C.J., Gong,J., Daugherty,B. and Rogers,N.
  TITLE     Leucine-rich repeat-containing g-protein coupled receptor-8
            molecules and uses thereof
  JOURNAL   Patent: WO 0214489-A 6 21-FEB-2002;
            Amgen, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2193
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             1. .2193
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD28588.1"
                     /db_xref="GI:19578157"
                     /translation="MIVFLVFKHLFSLRLITMFFLLHFIVLINVKDFALTQGSMITPS
                     CQKGYFPCGNLTKCLPRAFHCDGKDDCGNGADEENCGDTSGWATIFGTVHGNANSVAL
                     TQECFLKQYPQCCDCKETELECVNGDLKSVPMISNNVTLLSLKKNKIHSLPDKVFIKY
                     TKLKKIFLQHNCIRHISRKAFFGLCNLQILYLNHNCITTLRPGIFKDLHQLTWLILDD
                     NPITRISQRLFTGLNSLFFLSMVNNYLEALPKQMCAQMPQLNWVDLEGNRIKYLTNST
                     FLSCDSLTVLDLSSNTITELSPHLFKDLKLLQKLNLSSNPLMYLHKNQFESLKQLQSL
                     DLERIEIPNINTRMFQPMKNLSHIYFKNFRYCSYAPHVRICMPLTDGISSFEDLLANN
                     ILRIFVWVIAFITCFGNLFVIGMRSFIKAENTTHAMSIKILCCADCLMGVYLFFVGIF
                     DIKYRGQYQKYALLWMESVQCRLMGFLAMLSTEVSVLLLTYLTLEKFLVIVFPFSNIR
                     PGKRQTSVILICIWMAGFLIAVIPFWNKDYFGNFYGKNGVCFPLYYDQTEDIGSKGYS
                     LGIFLGVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDA
                     ICWIPVFVVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKLKQLL
                     HKHQRKSIFKIKKKSLSTSIVWIEDSSSLKLGVLNKITLGDSIMKPVS"
     sig_peptide     1. .108
BASE COUNT      626 a    456 c    411 g    700 t
ORIGIN

  Query Match           100.0%;  Score 321;  DB 6;  Length 2193;
  Best Local Similarity  100.0%;  Pred. No. 5.3e-71;
  Matches  321;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1714 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1773

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1774 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1833

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1834 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1893

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1894 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1953

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1954 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2013

Qy     301 AACTTTTTTAAGGACAAGTTG 321
          |||||||||||||||||||||
Db    2014 AACTTTTTTAAGGACAAGTTG 2034


RESULT 6
AX451562
LOCUS       AX451562               2214 bp    DNA     linear   PAT 03-JUL-2002
DEFINITION  Sequence 1 from Patent WO0226824.
ACCESSION   AX451562
```

```
VERSION     AX451562.1  GI:21698547
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Feder,J.N., Mintier,G., Ramanathan,C.S. and Hawken,D.R.
  TITLE     A novel human g-protein coupled receptor, hgprbmy5, expressed
            highly in brain and ovarian tissues
  JOURNAL   Patent: WO 0226824-A 1 04-APR-2002;
            Bristol-Myers Squibb Company (US)
FEATURES             Location/Qualifiers
     source          1. .2214
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      638 a    457 c    416 g    703 t
ORIGIN
```

Query Match 100.0%; Score 321; DB 6; Length 2214;
Best Local Similarity 100.0%; Pred. No. 5.2e-71;
Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
Db    1735 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1794

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
Db    1795 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1854

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
Db    1855 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1914

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
Db    1915 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1974

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
Db    1975 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2034

Qy     301 AACTTTTTTAAGGACAAGTTG 321
Db    2035 AACTTTTTTAAGGACAAGTTG 2055
```

```
RESULT 7
AX385027
LOCUS       AX385027               2265 bp    DNA     linear   PAT 19-MAR-2002
DEFINITION  Sequence 1 from Patent WO0214489.
ACCESSION   AX385027
VERSION     AX385027.1  GI:19578152
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Paszty,C.J., Gong,J., Daugherty,B. and Rogers,N.
  TITLE     Leucine-rich repeat-containing g-protein coupled receptor-8
            molecules and uses thereof
  JOURNAL   Patent: WO 0214489-A 1 21-FEB-2002;
            Amgen, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2265
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             1. .2265
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD28586.1"
                     /db_xref="GI:19578153"
                     /translation="MIVFLVFKHLFSLRLITMFFLLHFIVLINVKDFALTQGSMITPS
                     CQKGYFPCGNLTKCLPRAFHCDGKDDCGNGADEENCGDTSGWATIFGTVHGNANSVAL
                     TQECFLKQYPQCCDCKETELECVNGDLKSVPMISNNVTLLSLKKNKIHSLPDKVFIKY
                     TKLKKIFLQHNCIRHISRKAFFGLCNLQILYLNHNCITTLRPGIFKDLHQLTWLILDD
                     NPITRISQRLFTGLNSLFFLSMVNNYLEALPKQMCAQMPQLNWVDLEGNRIKYLTNST
                     FLSCDSLTVLFLPRNQIGFVPEKTFSSLKNLGELDLSSNTITELSPHLFKDLKLLQKL
                     NLSSNPLMYLHKNQFESLKQLQSLDLERIEIPNINTRMFQPMKNLSHIYFKNFRYCSY
                     APHVRICMPLTDGISSFEDLLANNILRIFVWVIAFITCFGNLFVIGMRSFIKAENTTH
                     AMSIKILCCADCLMGVYLFFVGIFDIKYRGQYQKYALLWMESVQCRLMGFLAMLSTEV
                     SVLLLTYLTLEKFLVIVFPFSNIRPGKRQTSVILICIWMAGFLIAVIPFWNKDYFGNF
                     YGKNGVCFPLYYDQTEDIGSKGYSLGIFLGVNLLAFLIIVFSYITMFCSIQKTALQTT
                     EVRNCFGREVAVANRFFFIVFSDAICWIPVFVVKILSLFRVEIPDTMTSWIVIFFLPV
                     NSALNPILYTLTTNFFKDKLKQLLHKHQRKSIFKIKKKSLSTSIVWIEDSSSLKLGVL
                     NKITLGDSIMKPVS"
     sig_peptide     1. .108
BASE COUNT      650 a    466 c    423 g    726 t
ORIGIN
```

Query Match 100.0%; Score 321; DB 6; Length 2265;
Best Local Similarity 100.0%; Pred. No. 5.2e-71;
Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
Db    1786 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1845

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
Db    1846 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1905

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
Db    1906 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1965

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
Db    1966 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 2025

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
Db    2026 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2085

Qy     301 AACTTTTTTAAGGACAAGTTG 321
Db    2086 AACTTTTTTAAGGACAAGTTG 2106
```

```
RESULT 8
AF453828
LOCUS       AF453828               2436 bp    mRNA    linear   PRI 04-FEB-2002
DEFINITION  Homo sapiens G protein-coupled receptor affecting testicular
            descent (GREAT) mRNA, complete cds.
ACCESSION   AF453828
VERSION     AF453828.1  GI:18483167
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2436)
  AUTHORS   Gorlov,I.P., Kamat,A., Jones,E., Lamb,D., Truong,A., Bogatcheva,N.,
            Bishop,C.E., McElreavey,K. and Agoulnik,A.I.
  TITLE     Mutations of the GREAT gene cause cryptorchidism
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2436)
  AUTHORS   Agoulnik,A.I.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-NOV-2001) Ob/Gyn, Baylor College of Medicine, 6550
            Fannin St., Su. 861, Houston, TX 77030, USA
FEATURES             Location/Qualifiers
     source          1. .2436
                     /organism="Homo sapiens"
```

```
                     /db_xref="taxon:9606"
                     /chromosome="13"
                     /map="13q12-q13"
     gene            1. .2436
                     /gene="GREAT"
     CDS             71. .2335
                     /gene="GREAT"
                     /codon_start=1
                     /product="G protein-coupled receptor affecting testicular
                     descent"
                     /protein_id="AAL73946.1"
                     /db_xref="GI:18483168"
                     /translation="MIVFLVFKHLFSLRLITMFFLLHFIVLINVKDFALTQGSMITPS
                     CQKGYFPCGNLTKCLPRAFHCDGKDDCGNGADEENCGDTSGWATIFGTVHGNANSVAL
                     TQECFLKQYPQCCDCKETELECVNGDLKSVPMISNNVTLLSLKKNKIHSLPDKVFIKY
                     TKLKKIFLQHNCIRHISRKAFFGLCNLQILYLNHNCITTLRPGIFKDLHQLTWLILDD
                     NPITRISQRLFTGLNSLFFLSMVNNYLEALPKQMCAQMPQLNWVDLEGNRIKYLTNST
                     FLSCDSLTVLFLPRNQIGFVPEKTFSSLKNLGELDLSSNTITELSPHLFKDLKLLQKL
                     NLSSNPLMYLHKNQFESLKQLQSLDLERIEIPNINTRMFQPMKNLSHIYFKNFRYCSY
                     APHVRICMPLTDGISSFEDLLANNILRIFVWVIAFITCFGNLFVIGMRSFIKAENTTH
                     AMSIKILCCADCLMGVYLFFVGIFDIKYRGQYQKYALLWMESVQCRLMGFLAMLSTEV
                     SVLLLTYLTLEKFLVIVFPFSNIRPGKRQTSVILICIWMAGFLIAVIPFWNKDYFGNF
                     YGKNGVCFPLYYDQTEDIGSKGYSLGIFLGVNLLAFLIIVFSYITMFCSIQKTALQTT
                     EVRNCFGREVAVANRFFFIVFSDAICWIPVFVVKILSLFRVEIPDTMTSWIVIFFLPV
                     NSALNPILYTLTTNFFKDKLKQLLHKHQRKSIFKIKKKSLSTSIVWIEDSSSLKLGVL
                     NKITLGDSIMKPVS"
BASE COUNT      703 a    503 c    456 g    774 t
ORIGIN

  Query Match           100.0%;  Score 321;  DB 9;  Length 2436;
  Best Local Similarity  100.0%;  Pred. No. 5.2e-71;
  Matches  321;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1856 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1915

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1916 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1975

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1976 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 2035

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2036 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 2095

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2096 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2155

Qy     301 AACTTTTTTAAGGACAAGTTG 321
          |||||||||||||||||||||
Db    2156 AACTTTTTTAAGGACAAGTTG 2176


RESULT 9
AF403384
LOCUS       AF403384               2838 bp    mRNA    linear   PRI 19-FEB-2002
DEFINITION  Homo sapiens LGR8 mRNA, complete cds.
ACCESSION   AF403384
VERSION     AF403384.2  GI:18702459
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2838)
  AUTHORS   Hsu,S.Y., Nakabayashi,K., Nishi,S., Kumagai,J., Kudo,M.,
            Sherwood,O.D. and Hsueh,A.J.
  TITLE     Activation of orphan receptors by the hormone relaxin
  JOURNAL   Science 295 (5555), 671-674 (2002)
  MEDLINE   21669315
   PUBMED   11809971
REFERENCE   2  (bases 1 to 2838)
  AUTHORS   Hsu,S.Y., Nakabayashi,K. and Bhalla,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUL-2001) GYN/OB, Stanford University, MSOB S385,
            Stanford, CA 94305, USA
REFERENCE   3  (bases 1 to 2838)
  AUTHORS   Hsueh,A.J.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-FEB-2002) Division of Reproductive Biology,
            Department of Gynecology and Obstetrics, Stanford University
            Medical Center, 300 Pasteur Drive, Room A-344, Stanford, CA
            94305-5317, USA
  REMARK    Sequence update by submitter
COMMENT     On Feb 19, 2002 this sequence version replaced gi:18419431.
FEATURES             Location/Qualifiers
     source          1. .2838
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11q13"
     CDS             107. .2371
                     /codon_start=1
                     /product="LGR8"
                     /protein_id="AAL69324.2"
                     /db_xref="GI:18702460"
                     /translation="MIVFLVFKHLFSLRLITMFFLLHFIVLINVKDFALTQGSMITPS
                     CQKGYFPCGNLTKCLPRAFHCDGKDDCGNGADEENCGDTSGWATIFGTVHGNANSVAL
                     TQECFLKQYPQCCDCKETELECVNGDLKSVPMISNNVTLLSLKKNKIHSLPDKVFIKY
                     TKLKKIFLQHNCIRHISRKAFFGLCNLQILYLNHNCITTLRPGIFKDLHQLTWLILDD
                     NPITRISQRLFTGLNSLFFLSMVNNYLEALPKQMCAQMPQLNWVDLEGNRIKYLTNST
                     FLSCDSLTVLFLPRNQIGFVPEKTFSSLKNLGELDLSSNTITELSPHLFKDLKLLQKL
                     NLSSNPLMYLHKNQFESLKQLQSLDLERIEIPNINTRMFQPMKNLSHIYFKNFRYCSY
                     APHVRICMPLTDGISSFEDLLANNILRIFVWVIAFITCFGNLFVIGMRSFIKAENTTH
                     AMSIKILCCADCLMGVYLFFVGIFDIKYRGQYQKYALLWMESVQCRLMGFLAMLSTEV
                     SVLLLTYLTLEKFLVIVFPFSNIRPGKRQTSVILICIWMAGFLIAVIPFWNKDYFGNF
                     YGKNGVCFPLYYDQTEDIGSKGYSLGIFLGVNLLAFLIIVFSYITMFCSIQKTALQTT
                     EVRNCFGREVAVANRFFFIVFSDAICWIPVFVVKILSLFRVEIPDTMTSWIVIFFLPV
                     NSALNPILYTLTTNFFKDKLKQLLHKHQRKSIFKIKKKSLSTSIVWIEDSSSLKLGVL
                     NKITLGDSIMKPVS"
BASE COUNT      817 a    616 c    536 g    869 t
ORIGIN

  Query Match           100.0%;  Score 321;  DB 9;  Length 2838;
  Best Local Similarity  100.0%;  Pred. No. 5.1e-71;
  Matches  321;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1892 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1951

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1952 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 2011

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2012 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 2071

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2072 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 2131

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2132 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2191

Qy     301 AACTTTTTTAAGGACAAGTTG 321
          |||||||||||||||||||||
Db    2192 AACTTTTTTAAGGACAAGTTG 2212
```

```
RESULT 10
AX385045
LOCUS       AX385045               2214 bp    DNA     linear   PAT 19-MAR-2002
DEFINITION  Sequence 19 from Patent WO0214489.
ACCESSION   AX385045
VERSION     AX385045.1  GI:19578166
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Paszty,C.J., Gong,J., Daugherty,B. and Rogers,N.
  TITLE     Leucine-rich repeat-containing g-protein coupled receptor-8
            molecules and uses thereof
  JOURNAL   Patent: WO 0214489-A 19 21-FEB-2002;
            Amgen, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2214
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     CDS             1. .2214
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD28593.1"
                     /db_xref="GI:19578167"
                     /translation="MWLLLHVILLTEVKDFALADSSMVAPLCPKGYFPCGNLTKCLPR
                     AFHCDGVDDCGNGADEDNCGDTSGWTTIFGTVHGNVNKVTLTQECFLSQYPQHCYCRE
                     NELECVKADLKAVPKVSSNVTLLSLKKNKIHRLPVKVFSRYTELRKIYLQHNCITHIS
                     RRAFLGLHNLQILYLSHNCITSLRPGIFKDLHQLAWLILDDNPITRISQKSFMGLNSL
                     FFLSMVGNRLEALPETLCAQMPQLNWVDLANNGIKYITNSTFLTCDSLTVLFLPRNQI
                     GFVPEKTFSSLKNLGELDLSSNMITKLPVHLFSDLHLLQKLNLSSNPLLYVHKNQFGS
                     LKQLQSLDLERIEIPNISTGMFQPMKNLSHIYLKTFRYCSYVPHVRICMPSTDGISSS
                     EDLLANGILRVSVWVIAFITCVGNFLVIAVRSLIKAENTTHAMSIKILCCADCLMGVY
                     LFSVGVFDIKYRGQYQKYALLWMESVPCRLLGFLATLSTEVSVLLLTFLTLEKFLVIV
                     FPFSNLRLGKRQTAVALASIWVVGFLIAAVPFTREDYFGNFYGKNGVCFPLHYDQAED
                     FGSRGYSLGIFLGVNLLAFLVIVISYVTMFCSIHKTALQTAEVRSHIGKEVAVANRFF
                     FIVFSDAICWIPVFVVKILSLLQVEIPGTITSWIVVFFLPVNSALNPILYTLTTSFFK
                     DKLKQLLHKHRRKPIFKVKKKSLSASIVWTDESSLKLGVLSKIALGDSIMKPVSP"
     sig_peptide     1. .57
BASE COUNT      554 a    548 c    492 g    620 t
ORIGIN

  Query Match             72.6%;  Score 233;  DB 6;  Length 2214;
  Best Local Similarity   82.9%;  Pred. No. 9.5e-49;
  Matches  266;  Conservative    0;  Mismatches   55;  Indels    0;  Gaps    0;

Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
           |||||||||||||||||||| ||| |||| ||| |||||||| | || |||||||| |||
Db    1735 GGTGTGAACTTGCTGGCTTTCCTCGTCATCGTGATTTCCTATGTCACCATGTTCTGCTCC 1794

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
           ||||| ||||| ||| | ||||| |||||||| ||||      | || |  |||||||||
Db    1795 ATTCATAAAACAGCCCTTCAGACTGCAGAAGTGAGGAGCCACATCGGGAAGGAGGTGGCT 1854

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
           |||||||| || |||||||||| |||||||||||||||||||||||||| ||||| |||
Db    1855 GTTGCAAACCGGTTCTTTTTTATCGTGTTCTCTGATGCCATCTGCTGGATCCCTGTGTTT 1914

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
           || ||||| ||||| || ||| | |  ||||| ||||||| |||||| ||||||||||
Db    1915 GTCGTTAAGATCCTGTCTCTCCTTCAAGTGGAGATACCAGGCACAATCACTTCCTGGATC 1974

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
           ||| |||||||||||||| || ||||| || || || || ||||||| ||||| || |||
Db    1975 GTGGTTTTTTTCCTTCCGGTGAACAGCGCCTTAAACCCCATCCTCTACACTCTGACGACC 2034

Qy     301 AACTTTTTTAAGGACAAGTTG 321
           |||||||||||||||||||||
Db    2035 TCCTTTTTTAAGGACAAGTTG 2055
```

```
RESULT 11
AF346501
LOCUS       AF346501               2539 bp    mRNA    linear   ROD 01-OCT-2001
DEFINITION  Mus musculus G protein coupled receptor affecting testicular
            descent (Great) mRNA, complete cds.
ACCESSION   AF346501
VERSION     AF346501.1  GI:15811372
KEYWORDS    .
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2539)
  AUTHORS   Overbeek,P.A., Gorlov,I.P., Sutherland,R.W., Houston,J.B.,
            Harrison,W.R., Boettger-Tong,H.L., Bishop,C.E. and Agoulnik,A.I.
  TITLE     A transgenic insertion causing cryptorchidism in mice
  JOURNAL   Genesis 30 (1), 26-35 (2001)
  MEDLINE   21250990
   PUBMED   11353515
REFERENCE   2  (bases 1 to 2539)
  AUTHORS   Agoulnik,A.I.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-FEB-2001) Ob/Gyn, Baylor College of Medicine, 6550
            Fannin Str. Su. 861, Houston, TX 77030, USA
FEATURES             Location/Qualifiers
     source          1. .2539
                     /organism="Mus musculus"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /chromosome="5"
     gene            1. .2539
                     /gene="Great"
     CDS             194. .2407
                     /gene="Great"
                     /note="GREAT"
                     /codon_start=1
                     /product="G protein coupled receptor affecting testicular
                     descent"
                     /protein_id="AAL08943.1"
                     /db_xref="GI:15811373"
                     /translation="MWLLLHVILLTEVKDFALADSSMVAPLCPKGYFPCGNLTKCLPR
                     AFHCDGVDDCGNGADEDNCGDTSGWTTIFGTVHGNVNKVTLTQECFLSQYPQHCYCRE
                     NELECVKADLKAVPKVSSNVTLLSLKKNKIHRLPVKVFSRYTELRKIYLQHNCITHIS
                     RRAFLGLHNLQILYLSHNCITSLRPGIFKDLHQLAWLILDDNPITRISQKSFMGLNSL
                     FFLPMVGNRLEALPETLCAQMPQLNWVDLANNGIKYITNSTFLTCDSLTVLFLPRNQI
                     GFVPEKTFSSLKNLGELDLSSNMITKLPVHLFSDLHLLQKLNLSSNPLLYVHKNQFGS
                     LKQLQSLDLERIEIPNISTGMFQPMKNLSHIYLKTFRYCSYVPHVRICMPSTDGISSS
                     EDLLANGILRVSVWVIAFITCVGNFLVIAVRSLIKAENTTHAMSIKILCCADCLMGVY
                     LFSVGVFDIKYRGQYQKYALLWMESVPCRLLGFLATLSTEVSVLLLTFLTLEKFLVIV
                     FPFSNLRLGKRQTAVALASIWVVGFLIAAVPFTREDYFGNFYGKNGVCFPLHYDQAED
                     FGSRGYSLGIFLGVNLLAFLVIVISYVTMFCSIHKTALQTAEVRSHIGKEVAVANRFF
                     FIVFSDAICWIPVFVVKILSLLQVEIPGTITSWIVVFFLPVNSALNPILYTLTTSFFK
                     DKLKQLLHKHRRKPIFKVKKKSLSASIVWTDESSLKLGVLSKIALGDSIMKPVSP"
BASE COUNT      639 a    641 c    574 g    685 t
ORIGIN

  Query Match             72.6%;  Score 233;  DB 10;  Length 2539;
  Best Local Similarity   82.9%;  Pred. No. 9.3e-49;
  Matches  266;  Conservative    0;  Mismatches   55;  Indels    0;  Gaps    0;

Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
           |||||||||||||||||||| ||| |||| ||| |||||||| | || |||||||| |||
Db    1928 GGTGTGAACTTGCTGGCTTTCCTCGTCATCGTGATTTCCTATGTCACCATGTTCTGCTCC 1987

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
           ||||| ||||| ||| | ||||| |||||||| ||||      | || |  |||||||||
Db    1988 ATTCATAAAACAGCCCTTCAGACTGCAGAAGTGAGGAGCCACATCGGGAAGGAGGTGGCT 2047

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
           |||||||| || |||||||||| |||||||||||||||||||||||||| ||||| |||
Db    2048 GTTGCAAACCGGTTCTTTTTTATCGTGTTCTCTGATGCCATCTGCTGGATCCCTGTGTTT 2107

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
```

```
          || ||||| ||||| || ||| | |  ||||| ||||||| |||||| ||||||||||
Db   2108 GTCGTTAAGATCCTGTCTCTCCTTCAAGTGGAGATACCAGGCACAATCACTTCCTGGATC 2167

Qy    241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
          ||| |||||||||||||| || ||||| || || || || |||||||| ||||| || |||
Db   2168 GTGGTTTTTTTCCTTCCGGTGAACAGCGCCTTAAACCCCATCCTCTACACTCTGACGACC 2227

Qy    301 AACTTTTTTAAGGACAAGTTG 321
          |||||||||||||||||||
Db   2228 TCCTTTTTTAAGGACAAGTTG 2248


RESULT 12
AX147820/c
LOCUS       AX147820               1018 bp    DNA     linear   PAT 08-JUN-2001
DEFINITION  Sequence 65 from Patent WO0136473.
ACCESSION   AX147820
VERSION     AX147820.1  GI:14346831
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1018)
  AUTHORS   Vogeli,G., Wood,L.S., Parodi,L.A., Hiebsch,R.R., Lind,P.,
            Slightom,J., Schellin,K.A., Kaytes,P.S., Bannigan,C.M., Ruff,V.,
            Sejlitz,T. and Huff,R.M.
  TITLE     Novel g protein-coupled receptors
  JOURNAL   Patent: WO 0136473-A 65 25-MAY-2001;
            PHARMACIA & UPJOHN COMPANY (US)
FEATURES             Location/Qualifiers
     source          1. .1018
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      332 a    179 c    189 g    318 t
ORIGIN

  Query Match             68.5%;  Score 220;  DB 6;  Length 1018;
  Best Local Similarity   100.0%;  Pred. No. 2e-45;
  Matches  220;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    539 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 480

Qy     61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    479 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 420

Qy    121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    419 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 360

Qy    181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAG 220
          ||||||||||||||||||||||||||||||||||||||||
Db    359 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAG 320


RESULT 13
AL138708
LOCUS       AL138708               170522 bp    DNA     linear   PRI 01-FEB-2001
DEFINITION  Human DNA sequence from clone RP11-432E15 on chromosome 13 Contains
            STSs, GSSs and a CpG island, complete sequence.
ACCESSION   AL138708
VERSION     AL138708.17  GI:11137618
KEYWORDS    HTG; CpG island.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 170522)
  AUTHORS   Phillimore,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Nov 10, 2000 this sequence version replaced gi:11121365.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 13, constructed by the Sanger Centre Chromosome 13
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr13
            This sequence is the entire insert of clone RP11-432E15 The true
            left end of clone RP11-154I23 is at 137612 in this sequence. The
            true right end of clone RP11-95N14 is at 13246 in this sequence.
            This sequence has been finished according to sequence map criteria
            as follows.  An attempt is made to resolve all sequencing problems,
            such as compressions and repeats, but not necessarily within known
            annotated repeat sequence elements.  Where the sequence is
            ambiguous, there is an annotation using the 'unsure' feature key.
            RP11-432E15 is from the library RPCI-11.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES             Location/Qualifiers
     source          1. .170522
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="13"
                     /clone="RP11-432E15"
                     /clone_lib="RPCI-11.2"
     misc_feature    19. .289
                     /note="match: GSS: Em:AQ549338"
     repeat_region   194. .299
                     /note="53 copies 2 mer tg 71% conserved"
     repeat_region   208. .299
                     /note="23 copies 4 mer tgtg 77% conserved"
     repeat_region   1456. .1750
                     /note="AluSx repeat: matches 3. .300 of consensus"
     repeat_region   2078. .2183
                     /note="MIR repeat: matches 6. .114 of consensus"
     misc_feature    2600. .3122
                     /note="match: GSS: Em:AQ529216"
     repeat_region   2646. .2940
                     /note="AluSg repeat: matches 1. .302 of consensus"
     repeat_region   3590. .3894
                     /note="AluSq repeat: matches 1. .296 of consensus"
     repeat_region   3926. .4462
                     /note="L2 repeat: matches 2233. .2750 of consensus"
     repeat_region   4658. .4805
                     /note="L2 repeat: matches 2355. .2503 of consensus"
     repeat_region   4911. .5224
                     /note="AluSq repeat: matches 1. .312 of consensus"
     repeat_region   5237. .5332
                     /note="L2 repeat: matches 2575. .2694 of consensus"
     repeat_region   5788. .6085
                     /note="AluJo repeat: matches 3. .301 of consensus"
     misc_feature    7355. .7747
                     /note="sequence from overlapping clone bA159O9
                     (AL136106).Assembly confirmed by restriction digest."
     repeat_region   7388. .7487
                     /note="50 copies 2 mer at 63% conserved"
     repeat_region   7396. .7487
                     /note="23 copies 4 mer atat 66% conserved"
     repeat_region   7884. .7915
                     /note="MER20 repeat: matches 30. .62 of consensus"
```

```
repeat_region   7916. .8219
                /note="AluJb repeat: matches 1. .302 of consensus"
repeat_region   8220. .8376
                /note="MER20 repeat: matches 62. .210 of consensus"
misc_feature    8975. .8985
                /note="sequence from overlapping clone
                bA159O9(AL136106).Assembly confirmed by restriction
                digest."
repeat_region   9078. .9121
                /note="11 copies 4 mer caca 100% conserved"
misc_feature    complement(9603. .10104)
                /note="match: GSS: Em:AQ569313"
misc_feature    11234. .11331
                /note="sequence from overlapping clone bA159O9
                (AL136106).Assembly confirmed by restriction digest."
misc_feature    complement(11953. .12400)
                /note="match: GSS: Em:AQ486928"
repeat_region   12477. .12601
                /note="MIR repeat: matches 15. .138 of consensus"
misc_feature    12713. .12980
                /note="sequence from overlapping clone bA159O9
                (AL136106).Assembly confirmed by restriction digest."
repeat_region   13272. .13417
                /note="MIR repeat: matches 38. .188 of consensus"
misc_feature    complement(13670. .13919)
                /note="match: GSS: Em:AQ021664"
misc_feature    complement(13740. .14130)
                /note="match: GSS: Em:AQ032630"
misc_feature    13771. .14227
                /note="sequence from overlapping clone bA159O9
                (AL136106).Assembly confirmed by restriction digest."
repeat_region   13864. .14087
                /note="L2 repeat: matches 2515. .2750 of consensus"
repeat_region   16061. .16342
                /note="MER1B repeat: matches 1. .244 of consensus"
repeat_region   16343. .16655
                /note="AluSx repeat: matches 1. .312 of consensus"
repeat_region   16656. .16680
                /note="MER1B repeat: matches 244. .305 of consensus"
misc_feature    complement(17019. .17275)
                /note="match: GSS: Em:AQ452676"
misc_feature    17070. .17578
                /note="match: GSS: Em:AQ511484"
misc_feature    17087. .17515
                /note="match: GSS: Em:AQ817416"
misc_feature    17278. .17794
                /note="match: GSS: Em:AQ358157"
repeat_region   18301. .18381
                /note="L1MB2 repeat: matches 6082. .6166 of consensus"
misc_feature    complement(18563. .19010)
                /note="match: GSS: Em:AQ242897"
repeat_region   20403. .20559
                /note="MER20 repeat: matches 1. .153 of consensus"
repeat_region   20870. .21069
                /note="100 copies 2 mer tt 69% conserved"
repeat_region   20937. .21076
                /note="35 copies 4 mer tttc 72% conserved"
repeat_region   21080. .21392
                /note="AluJb repeat: matches 1. .310 of consensus"
repeat_region   21395. .21538
                /note="MIR repeat: matches 94. .245 of consensus"
repeat_region   23716. .23960
                /note="L2 repeat: matches 2381. .2617 of consensus"
repeat_region   23961. .24213
                /note="AluSx repeat: matches 47. .299 of consensus"
repeat_region   24214. .24342
                /note="L2 repeat: matches 2617. .2747 of consensus"
repeat_region   25636. .25785
                /note="MER5A repeat: matches 32. .189 of consensus"
repeat_region   25638. .25810
                /note="MER5A repeat: matches 17. .189 of consensus"
repeat_region   26013. .26227
                /note="MIR repeat: matches 10. .256 of consensus"
repeat_region   28426. .28484
                /note="L2 repeat: matches 2641. .2700 of consensus"
repeat_region   28639. .28960
                /note="AluSq repeat: matches 1. .310 of consensus"
repeat_region   29689. .29861
                /note="L2 repeat: matches 2006. .2186 of consensus"
misc_feature    complement(29720. .30187)
                /note="match: GSS: Em:AQ317477"
repeat_region   29901. .30075
                /note="MER5A repeat: matches 1. .184 of consensus"
misc_feature    complement(30119. .30414)
                /note="match: GSS: Em:AQ230143"
repeat_region   32462. .32774
                /note="AluSx repeat: matches 1. .312 of consensus"
repeat_region   32807. .32877
                /note="L2 repeat: matches 2635. .2707 of consensus"
repeat_region   35365. .35806
                /note="L2 repeat: matches 2248. .2742 of consensus"
misc_feature    complement(35852. .36413)
                /note="match: GSS: Em:AQ284653
                match: STS: Em:G55301"
misc_feature    complement(36043. .36433)
                /note="match: GSS: Em:AQ816559"
repeat_region   36420. .36469
                /note="MIR repeat: matches 96. .148 of consensus"
misc_feature    36429. .36961
                /note="match: GSS: Em:AQ394779"
misc_feature    36452. .36845
                /note="match: GSS: Em:AQ212353"
repeat_region   39203. .39242
                /note="20 copies 2 mer ca 100% conserved"
repeat_region   39250. .39648
                /note="L2 repeat: matches 2307. .2710 of consensus"
repeat_region   39703. .39750
                /note="24 copies 2 mer ca 75% conserved"
repeat_region   39859. .39976
                /note="L1MC3 repeat: matches 7621. .7739 of consensus"
repeat_region   39988. .40283
                /note="AluSx repeat: matches 1. .296 of consensus"
repeat_region   40873. .41144
                /note="AluY repeat: matches 39. .310 of consensus"
```

```
  Query Match             68.5%;  Score 220;  DB 9;  Length 170522;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-45;
  Matches  220;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   42034 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 42093

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   42094 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 42153

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   42154 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 42213

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAG 220
          ||||||||||||||||||||||||||||||||||||||||
Db   42214 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAG 42253
```

```
RESULT 14
AX128507
LOCUS       AX128507               1191 bp    DNA     linear   PAT 15-MAY-2001
DEFINITION  Sequence 15 from Patent WO0131014.
ACCESSION   AX128507
VERSION     AX128507.1  GI:14134974
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1191)
  AUTHORS   Vogeli,G., Wood,L.S. and Merchant,K.
  TITLE     G protein-coupled receptors expressed in brain
  JOURNAL   Patent: WO 0131014-A 15 03-MAY-2001;
            PHARMACIA & UPJOHN COMPANY (US)
FEATURES             Location/Qualifiers
     source          1. .1191
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             1. .1191
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC38938.1"
                     /db_xref="GI:14134975"
                     /translation="MFRPLVNLSHIYFKKFQYCGYAPHVRSCKPNTDGISSLENLLAS
                     IIQRVFVWVVSAVTCFGNIFVICMRPYIRSENKLYAMSIISLCCADCLMGIYLFVIGG
                     FDLKFRGEYNKHAQLWMESTHCQLVGSLAILSTEVSVLLLTFLTLEKYICIVYPFRCV
                     RPGKCRTITVLILIWITGFIVAFIPLSNKEFFKNYYGTNGVCFPLHSEDTESIGAQIY
                     SVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVFTD
                     ALCWIPIFVVKFLSLLQVEIPGTITSWVVIFILPINSALNPILYTLTTRPFKEMIHRF
                     WYNYRQRKSMDSKGQKTYAPSFIWVEMWPLQEMPPELMKPDLFTYPCEMSLISQSTRL
                     NSYS"
BASE COUNT      340 a    229 c    226 g    396 t
ORIGIN
```

Query Match 51.8%; Score 166.4; DB 6; Length 1191;
Best Local Similarity 70.0%; Pred. No. 7.1e-32;
Matches 224; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

```
Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
Db     673 GGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGT 732

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
Db     733 GTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATC 792

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
Db     793 CTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTT 852

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
Db     853 GTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTA 912

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
Db     913 GTGATTTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACA 972

Qy     301 AACTTTTTTAAGGACAAGTT 320
Db     973 AGACCATTTAAAGAAATGAT 992
```

```
RESULT 15
AF190500
LOCUS       AF190500               2274 bp    mRNA    linear   PRI 10-AUG-2001
DEFINITION  Homo sapiens leucine-rich repeat-containing G protein-coupled
            receptor 7 (LGR7) mRNA, complete cds.
ACCESSION   AF190500
VERSION     AF190500.1  GI:10441729
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2274)
  AUTHORS   Hsu,S.Y., Kudo,M., Chen,T., Nakabayashi,K., Bhalla,A., van der
            Spek,P.J., van Duin,M. and Hsueh,A.J.
  TITLE     The three subfamilies of leucine-rich repeat-containing G
            protein-coupled receptors (LGR): identification of LGR6 and LGR7
            and the signaling mechanism for LGR7
  JOURNAL   Mol. Endocrinol. 14 (8), 1257-1271 (2000)
  MEDLINE   20388592
   PUBMED   10935549
REFERENCE   2  (bases 1 to 2274)
  AUTHORS   Hsu,S.Y. and Hsueh,A.J.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-SEP-1999) GYN/OB, Stanford University, MSOB S385,
            Stanford, CA 94305, USA
FEATURES             Location/Qualifiers
     source          1. .2274
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            1. .2274
                     /gene="LGR7"
     CDS             1. .2274
                     /gene="LGR7"
                     /codon_start=1
                     /product="leucine-rich repeat-containing G protein-coupled
                     receptor 7"
                     /protein_id="AAG17167.1"
                     /db_xref="GI:10441730"
                     /translation="MTSGSVFFYILIFGKYFSHGGGQDVKCSLGYFPCGNITKCLPQL
                     LHCNGVDDCGNQADEDNCGDNNGWSMQFDKYFASYYKMTSQYPFEAETPECLVGSVPV
                     QCLCQGLELDCDETNLRAVPSVSSNVTAMSLQWNLIRKLPPDCFKNYHDLQKLYLQNN
                     KITSISIYAFRGLNSLTKLYLSHNRITFLKPGVFEDLHRLEWLIIEDNHLSRISPPTF
                     YGLNSLILLVLMNNVLTRLPDKPLCQHMPRLHWLDLEGNHIHNLRNLTFISCSNLTVL
                     VMRKNKINHLNENTFAPLQKLDELDLGSNKIENLPPLIFKDLKELSQLNLSYNPIQKI
                     QANQFDYLVKLKSLSLEGIEISNIQQRMFRPLMNLSHIYFKKFQYCGYAPHVRSCKPN
                     TDGISSLENLLASIIQRVFVWVVSAVTCFGNIFVICMRPYIRSENKLYAMSIISLCCA
                     DCLMGIYLFVIGGFDLKFRGEYNKHAQLWMESTHCQLVGSLAILSTEVSVLLLTFLTL
                     EKYICIVYPFRCVRPGKCRTITVLILIWITGFIVAFIPLSNKEFFKNYYGTNGVCFPL
                     HSEDTESIGAQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEM
                     ILAKRFFFIVFTDALCWIPIFVVKFLSLLQVEIPGTITSWVVIFILPINSALNPILYT
                     LTTRPFKEMIHRFWYNYRQRKSMDSKGQKTYAPSFIWVEMWPLQEMPPELMKPDLFTY
                     PCEMSLISQSTRLNSYS"
BASE COUNT      669 a    460 c    421 g    724 t
ORIGIN
```

Query Match 51.8%; Score 166.4; DB 9; Length 2274;
Best Local Similarity 70.0%; Pred. No. 6.6e-32;
Matches 224; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

```
Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
Db    1756 GGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGT 1815

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
Db    1816 GTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATC 1875

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
Db    1876 CTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTT 1935

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
Db    1936 GTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTA 1995

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
Db    1996 GTGATTTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACA 2055

Qy     301 AACTTTTTTAAGGACAAGTT 320
Db    2056 AGACCATTTAAAGAAATGAT 2075
```

Search completed: May 26, 2003, 19:35:52
Job time : 1994 secs

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 17:58:02 ; Search time 233 Seconds
(without alignments)
3102.538 Million cell updates/sec

Title: US-09-930-312-1
Perfect score: 321
Sequence: 1 ggtgtgaacttgctggcttt.....acttttttaaggacaagttg 321

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	321	100.0	321	24	AAD32027	Human novel G-prot
	2	321	100.0	1065	24	ABL40194	Human G protein-co
	3	321	100.0	1068	22	AAS07943	Human cDNA encodin
	4	321	100.0	1473	24	ABL40192	Human G protein-co
	5	321	100.0	1545	24	ABL40191	Human G protein-co
	6	321	100.0	1830	24	ABL40188	Human G protein-co
	7	321	100.0	2142	24	ABK51947	cDNA encoding huma
	8	321	100.0	2190	24	ABL40197	Human G protein-co
	9	321	100.0	2214	24	ABK51944	cDNA encoding huma
	10	321	100.0	2262	24	ABL40196	Human G protein-co
c	11	220	68.5	1018	22	AAH51001	Human nGPCR57 codi
	12	166.4	51.8	530	22	AAI99584	Human expressed po
	13	166.4	51.8	530	22	AAL35644	Human musculoskele
	14	166.4	51.8	530	22	ABA06471	Human cDNA SEQ ID
	15	166.4	51.8	530	22	AAS28950	cDNA encoding for
	16	166.4	51.8	530	22	AAS29573	Human endocrine po
	17	166.4	51.8	530	22	AAS30187	DNA encoding rena
	18	166.4	51.8	530	22	AAS34845	cDNA encoding nove
	19	166.4	51.8	530	23	ABK43875	DNA encoding novel
	20	166.4	51.8	1089	22	AAI99557	Human expressed po
c	21	166.4	51.8	1162	23	ABK43573	DNA encoding novel
	22	166.4	51.8	1191	22	AAD06507	Human CON222 G pro
	23	166.4	51.8	1804	21	AAZ90524	Human GPCR protein
	24	166.4	51.8	2467	20	AAZ25345	Human LGR7 long fo
	25	166.4	51.8	3584	20	AAZ25346	Human LGR7 short f
	26	165.4	51.5	420	21	AAC75735	Human ORFX ORF1290
	27	156.8	48.8	474	22	AAF28059	Human HGRL101 G-pr
	28	67.2	20.9	1015	22	AAS57085	cDNA encoding Dros
	29	67.2	20.9	1015	23	ABL07293	Drosophila melanog
	30	53.2	16.6	1080	22	AAS57089	cDNA encoding Dros
	31	53.2	16.6	1080	23	ABL08799	Drosophila melanog
	32	53.2	16.6	3187	22	AAS57088	DNA encoding Droso
	33	53.2	16.6	3187	23	ABL08798	Drosophila melanog
	34	53.2	16.6	3494	23	ABL08218	Drosophila melanog
	35	46.6	14.5	636	21	AAA44932	Human secreted exp
	36	39.2	12.2	3574	22	AAS57084	DNA encoding Droso
	37	39.2	12.2	3574	23	ABL07292	Drosophila melanog
	38	39.2	12.2	3578	23	ABL03716	Drosophila melanog
	39	36.6	11.4	14253	24	ABL33494	Human immune syste
	40	36	11.2	6341	24	ABL33645	Human immune syste
	41	36	11.2	6624	24	ABL32043	Human immune syste
	42	35.2	11.0	201	21	AAC31368	Human secreted pro
c	43	35.2	11.0	535	23	ABV52410	Human prostate exp
	44	35.2	11.0	26464	22	AAK81246	Human immune/haema
c	45	35.2	11.0	26747	24	AAI67784	Nucleotide sequenc

ALIGNMENTS

RESULT 1
AAD32027
ID AAD32027 standard; DNA; 321 BP.
XX
AC AAD32027;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human novel G-protein coupled receptor (nGPCR)-1079 DNA.
XX
KW Human; novel G-protein coupled receptor; nGPCR-1079; viral infection;
KW gene; therapy; human immunodeficiency virus; HIV; pain; migraine;
KW central nervous system disorder; stroke; manic depression; obesity;
KW metabolic disorder; anorexia; cardiovascular disorder; type 2 diabetes;
KW cancer; myocardial infarction; hypotension; degenerative disorder;
KW Parkinson's disease; Alzheimer's disease; neurological disorder;
KW schizophrenia; anxiety; inflammatory condition; rheumatoid arthritis;
KW thyroid disorder; autoimmune disorder; hormonal disorder; renal failure;
KW psoriasis; movement disorder; analgesic; cytostatic; neuroprotective;
KW anorectic; hypotensive; hypertensive; tranquilliser; anticonvulsant;
KW metabolic; neuroleptic; thrombolytic; cardiant; immunosuppressive; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..321
FT /*tag= a
FT /product= "Human nGPCR-1079 protein"
FT /note= "CDS does not include start and stop codon"
FT /partial
XX
PN WO200214496-A2.

XX
PD 21-FEB-2002.
XX
PF 15-AUG-2001; 2001WO-US25513.
XX
PR 15-AUG-2000; 2000US-225262P.
XX
PA (PHAA) PHARMACIA & UPJOHN CO.
XX
PI Lind P;
XX
DR WPI; 2002-269192/31.
DR P-PSDB; AAE20148.
XX
PT An isolated nucleic acid molecule encoding novel G-protein coupled
PT receptor polypeptide which is useful for treating obesity, diabetes,
PT Parkinson's disease, manic depression, migraine, rheumatoid arthritis
PT -
XX
PS Claim 4; Page 60; 93pp; English.
XX
CC The patent discloses novel G-protein coupled receptor (nGPCR)-1079 genes
CC and their corresponding proteins. Sequences of the invention are useful
CC for treating diseases such as viral infections caused by human immuno-
CC deficiency virus (HIV)-1 or 2, central nervous system (CNS) disorders
CC (e.g. pain, including migraine, stroke, manic depression), metabolic
CC disorders (e.g. obesity and anorexia), cancers, cardiovascular disorders
CC (e.g. type 2 diabetes, myocardial infarction, hypotension), degenerative
CC disorders (e.g. Parkinson's disease, Alzheimer's disease), neurological
CC disorders (e.g. schizophrenia and anxiety), inflammatory conditions,
CC rheumatoid arthritis, thyroid disorders, autoimmune disorders, hormonal
CC disorders, renal failure, psoriasis and movement disorders. The present
CC sequence is a DNA encoding human nGPCR-1079 protein.
XX
SQ Sequence 321 BP; 76 A; 67 C; 59 G; 119 T; 0 other;

Query Match 100.0%; Score 321; DB 24; Length 321;
Best Local Similarity 100.0%; Pred. No. 1.2e-86;
Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60

Qy     61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120

Qy    121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180

Qy    181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240

Qy    241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300

Qy    301 AACTTTTTTAAGGACAAGTTG 321
          |||||||||||||||||||||
Db    301 AACTTTTTTAAGGACAAGTTG 321
```

RESULT 2
ABL40194
ID ABL40194 standard; cDNA; 1065 BP.
XX
AC ABL40194;
XX
DT 23-MAY-2002 (first entry)
XX
DE Human G protein-coupled receptor TGR17-4 encoding cDNA SEQ ID NO:11.
XX
KW Human; G protein-coupled receptor; TGR17-4; nootropic; antiinflammatory;
KW vasotropic; immunomodulator; cytostatic; gene therapy; protein therapy;
KW neurological; inflammatory; circulatory; degenerative; immune system;
KW digestive disease; cancer; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1065
FT /*tag= a
FT /partial
FT /product= "TGR17-4"
FT /note= "no stop codon given"
XX
PN WO200204640-A1.
XX
PD 17-JAN-2002.
XX
PF 06-JUL-2001; 2001WO-JP05878.
XX
PR 07-JUL-2000; 2000JP-0211989.
PR 18-DEC-2000; 2000JP-0383794.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Moriya T, Ito T, Shintani Y, Miyajima N;
XX
DR WPI; 2002-179706/23.
DR P-PSDB; ABB06253.
XX
PT G-protein coupled receptor protein TGR17 of human origin and DNA
PT encoding it for diagnosis and treatment of cancer and circulatory and
PT other diseases associated with its expression -
XX
PS Claim 7; Page 125-126; 145pp; Japanese.
XX
CC The present invention describes a human guanine nucleotide binding
CC protein (G protein) coupled receptor protein designated TGR17, which has
CC 5 variant forms (TGR17-1 to TGR17-5) of the receptor protein having
CC additional residues at the N-terminal. The TGR17 proteins have nootropic,
CC antiinflammatory, vasotropic, immunomodulator and cytostatic activities.
CC The TGR17 polynucleotides and protein can be used in gene therapy and
CC protein therapy. G-protein coupled receptor proteins are cell membrane
CC proteins mediating the cellular response to a large variety of signalling
CC molecules. The TGR17 polynucleotides and proteins can be used in the
CC diagnosis, treatment and prevention of diseases including neurological,
CC inflammatory, circulatory, degenerative, immune system and digestive
CC diseases and cancer. The present sequence encodes human TGR17-4 from the
CC present invention.
XX
SQ Sequence 1065 BP; 277 A; 214 C; 213 G; 361 T; 0 other;

Query Match 100.0%; Score 321; DB 24; Length 1065;
Best Local Similarity 100.0%; Pred. No. 1.9e-86;
Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    589 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 648

Qy     61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 708

Qy    121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 768

Qy    181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 828
```

```
Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     829 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 888

Qy     301 AACTTTTTTAAGGACAAGTTG 321
           |||||||||||||||||||||
Db     889 AACTTTTTTAAGGACAAGTTG 909


RESULT 3
AAS07943
ID   AAS07943 standard; cDNA; 1068 BP.
XX
AC   AAS07943;
XX
DT   23-OCT-2001  (first entry)
XX
DE   Human cDNA encoding G-protein coupled receptor, hRUP16.
XX
KW   Human; G-protein coupled receptor; GPCR; hRUP16; agonist;
KW   inverse agonist; lung cancer; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1068
FT                   /*tag=  a
FT                   /product= "hRUP16"
XX
PN   WO200136471-A2.
XX
PD   25-MAY-2001.
XX
PF   16-NOV-2000; 2000WO-US31509.
XX
PR   17-NOV-1999;   99US-0166088.
PR   17-NOV-1999;   99US-0166099.
PR   17-NOV-1999;   99US-0166369.
PR   23-DEC-1999;   99US-0171900.
PR   23-DEC-1999;   99US-0171901.
PR   23-DEC-1999;   99US-0171902.
PR   11-FEB-2000; 2000US-0181749.
PR   14-MAR-2000; 2000US-0189258.
PR   14-MAR-2000; 2000US-0189259.
PR   10-APR-2000; 2000US-0195898.
PR   10-APR-2000; 2000US-0195899.
PR   10-APR-2000; 2000US-0196078.
PR   28-APR-2000; 2000US-0200419.
PR   12-MAY-2000; 2000US-0203630.
PR   12-JUN-2000; 2000US-0210741.
PR   12-JUN-2000; 2000US-0210982.
PR   21-AUG-2000; 2000US-0226760.
PR   26-SEP-2000; 2000US-0235418.
PR   26-SEP-2000; 2000US-0235779.
PR   20-OCT-2000; 2000US-0242332.
PR   20-OCT-2000; 2000US-0242343.
XX
PA   (AREN-) ARENA PHARM INC.
XX
PI   Chen R,  Dang HT,  Lowitz KP;
XX
DR   WPI; 2001-355616/37.
DR   P-PSDB; AAU04370.
XX
PT   Endogenous and non-endogenous versions of human G-protein coupled
PT   receptors for direct identification of candidate compounds as agonists,
PT   inverse agonists or partial agonists for use as therapeutic agents -
XX
PS   Claim 35; Page 104-105; 159pp; English.
XX
CC   The sequence encodes a human G-protein coupled receptor (GPCR),
CC   hRUP16 The endogenous and non-endogenous, constitutively activated
CC   versions of human G-protein coupled receptors (GPCR), are useful for
CC   direct identification of candidate compounds as receptor agonists,
CC   inverse agonists or partial agonists having applicability as therapeutic
CC   agents for treating diseases related to GPCR, e.g. lung cancer.
CC   Non-endogenous version of human GPCRs are also utilized in research
CC   settings and in vitro and in vivo system, incorporating GPCRs can be
CC   utilised to elucidate and understand the roles these receptors
CC   play in the human condition, both normal and diseased.
XX
SQ   Sequence 1068 BP; 278 A; 215 C; 214 G; 361 T; 0 other;

  Query Match          100.0%;  Score 321;  DB 22;  Length 1068;
  Best Local Similarity  100.0%;  Pred. No. 1.9e-86;
  Matches  321;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     589 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 648

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     649 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 708

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     709 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 768

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     769 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 828

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     829 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 888

Qy     301 AACTTTTTTAAGGACAAGTTG 321
           |||||||||||||||||||||
Db     889 AACTTTTTTAAGGACAAGTTG 909


RESULT 4
ABL40192
ID   ABL40192 standard; cDNA; 1473 BP.
XX
AC   ABL40192;
XX
DT   23-MAY-2002  (first entry)
XX
DE   Human G protein-coupled receptor TGR17-3 encoding cDNA SEQ ID NO:8.
XX
KW   Human; G protein-coupled receptor; TGR17-3; nootropic; antiinflammatory;
KW   vasotropic; immunomodulator; cytostatic; gene therapy; protein therapy;
KW   neurological; inflammatory; circulatory; degenerative; immune system;
KW   digestive disease; cancer; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1473
FT                   /*tag=  a
FT                   /partial
FT                   /product= "TGR17-3"
FT                   /note= "no stop codon given"
XX
PN   WO200204640-A1.
XX
PD   17-JAN-2002.
XX
PF   06-JUL-2001; 2001WO-JP05878.
XX
PR   07-JUL-2000; 2000JP-0211989.
PR   18-DEC-2000; 2000JP-0383794.
XX
```

PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Moriya T, Ito T, Shintani Y, Miyajima N;
XX
DR WPI; 2002-179706/23.
DR P-PSDB; ABB06252.
XX
PT G-protein coupled receptor protein TGR17 of human origin and DNA
PT encoding it for diagnosis and treatment of cancer and circulatory and
PT other diseases associated with its expression -
XX
PS Claim 7; Page 122-123; 145pp; Japanese.
XX
CC The present invention describes a human guanine nucleotide binding
CC protein (G protein) coupled receptor protein designated TGR17, which has
CC 5 variant forms (TGR17-1 to TGR17-5) of the receptor protein having
CC additional residues at the N-terminal. The TGR17 proteins have nootropic,
CC antiinflammatory, vasotropic, immunomodulator and cytostatic activities.
CC The TGR17 polynucleotides and protein can be used in gene therapy and
CC protein therapy. G-protein coupled receptor proteins are cell membrane
CC proteins mediating the cellular response to a large variety of signalling
CC molecules. The TGR17 polynucleotides and proteins can be used in the
CC diagnosis, treatment and prevention of diseases including neurological,
CC inflammatory, circulatory, degenerative, immune system and digestive
CC diseases and cancer. The present sequence encodes human TGR17-3 from the
CC present invention.
XX
SQ Sequence 1473 BP; 405 A; 310 C; 278 G; 480 T; 0 other;

Query Match 100.0%; Score 321; DB 24; Length 1473;
Best Local Similarity 100.0%; Pred. No. 2.1e-86;
Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
Db 997 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1056

Qy 61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
Db 1057 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1116

Qy 121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
Db 1117 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1176

Qy 181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
Db 1177 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1236

Qy 241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
Db 1237 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 1296

Qy 301 AACTTTTTTAAGGACAAGTTG 321
Db 1297 AACTTTTTTAAGGACAAGTTG 1317

RESULT 5
ABL40191
ID ABL40191 standard; cDNA; 1545 BP.
XX
AC ABL40191;
XX
DT 23-MAY-2002 (first entry)
XX
DE Human G protein-coupled receptor TGR17-2 encoding cDNA SEQ ID NO:6.
XX
KW Human; G protein-coupled receptor; TGR17-2; nootropic; antiinflammatory;
KW vasotropic; immunomodulator; cytostatic; gene therapy; protein therapy;
KW neurological; inflammatory; circulatory; degenerative; immune system;
KW digestive disease; cancer; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1545
FT /*tag= a
FT /partial
FT /product= "TGR17-2"
FT /note= "no stop codon given"
XX
PN WO200204640-A1.
XX
PD 17-JAN-2002.
XX
PF 06-JUL-2001; 2001WO-JP05878.
XX
PR 07-JUL-2000; 2000JP-0211989.
PR 18-DEC-2000; 2000JP-0383794.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Moriya T, Ito T, Shintani Y, Miyajima N;
XX
DR WPI; 2002-179706/23.
DR P-PSDB; ABB06251.
XX
PT G-protein coupled receptor protein TGR17 of human origin and DNA
PT encoding it for diagnosis and treatment of cancer and circulatory and
PT other diseases associated with its expression -
XX
PS Claim 7; Page 118-119; 145pp; Japanese.
XX
CC The present invention describes a human guanine nucleotide binding
CC protein (G protein) coupled receptor protein designated TGR17, which has
CC 5 variant forms (TGR17-1 to TGR17-5) of the receptor protein having
CC additional residues at the N-terminal. The TGR17 proteins have nootropic,
CC antiinflammatory, vasotropic, immunomodulator and cytostatic activities.
CC The TGR17 polynucleotides and protein can be used in gene therapy and
CC protein therapy. G-protein coupled receptor proteins are cell membrane
CC proteins mediating the cellular response to a large variety of signalling
CC molecules. The TGR17 polynucleotides and proteins can be used in the
CC diagnosis, treatment and prevention of diseases including neurological,
CC inflammatory, circulatory, degenerative, immune system and digestive
CC diseases and cancer. The present sequence encodes human TGR17-2 from the
CC present invention.
XX
SQ Sequence 1545 BP; 429 A; 320 C; 290 G; 506 T; 0 other;

Query Match 100.0%; Score 321; DB 24; Length 1545;
Best Local Similarity 100.0%; Pred. No. 2.1e-86;
Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
Db 1069 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1128

Qy 61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
Db 1129 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1188

Qy 121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
Db 1189 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1248

Qy 181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
Db 1249 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1308

Qy 241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
Db 1309 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 1368

Qy 301 AACTTTTTTAAGGACAAGTTG 321

```
Db   1369 AACTTTTTTAAGGACAAGTTG 1389


RESULT 6
ABL40188
ID   ABL40188 standard; cDNA; 1830 BP.
XX
AC   ABL40188;
XX
DT   23-MAY-2002  (first entry)
XX
DE   Human G protein-coupled receptor TGR17-1 encoding cDNA SEQ ID NO:2.
XX
KW   Human; G protein-coupled receptor; TGR17-1; nootropic; antiinflammatory;
KW   vasotropic; immunomodulator; cytostatic; gene therapy; protein therapy;
KW   neurological; inflammatory; circulatory; degenerative; immune system;
KW   digestive disease; cancer; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1830
FT                   /*tag= a
FT                   /partial
FT                   /product= "TGR17-1"
FT                   /note= "no stop codon given"
XX
PN   WO200204640-A1.
XX
PD   17-JAN-2002.
XX
PF   06-JUL-2001; 2001WO-JP05878.
XX
PR   07-JUL-2000; 2000JP-0211989.
PR   18-DEC-2000; 2000JP-0383794.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
PI   Moriya T,  Ito T,  Shintani Y,  Miyajima N;
XX
DR   WPI; 2002-179706/23.
DR   P-PSDB; ABB06250.
XX
PT   G-protein coupled receptor protein TGR17 of human origin and DNA
PT   encoding it for diagnosis and treatment of cancer and circulatory and
PT   other diseases associated with its expression -
XX
PS   Claim 7; Page 114-115; 145pp; Japanese.
XX
CC   The present invention describes a human guanine nucleotide binding
CC   protein (G protein) coupled receptor protein designated TGR17, which has
CC   5 variant forms (TGR17-1 to TGR17-5) of the receptor protein having
CC   additional residues at the N-terminal. The TGR17 proteins have nootropic,
CC   antiinflammatory, vasotropic, immunomodulator and cytostatic activities.
CC   The TGR17 polynucleotides and protein can be used in gene therapy and
CC   protein therapy. G-protein coupled receptor proteins are cell membrane
CC   proteins mediating the cellular response to a large variety of signalling
CC   molecules. The TGR17 polynucleotides and proteins can be used in the
CC   diagnosis, treatment and prevention of diseases including neurological,
CC   inflammatory, circulatory, degenerative, immune system and digestive
CC   diseases and cancer. The present sequence encodes human TGR17-1 from the
CC   present invention.
XX
SQ   Sequence 1830 BP; 516 A; 377 C; 351 G; 586 T; 0 other;

  Query Match             100.0%;  Score 321;  DB 24;  Length 1830;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-86;
  Matches  321;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1354 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1413

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1414 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1473

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1474 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1533

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1534 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1593

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1594 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 1653

Qy     301 AACTTTTTTAAGGACAAGTTG 321
           |||||||||||||||||||||
Db    1654 AACTTTTTTAAGGACAAGTTG 1674


RESULT 7
ABK51947
ID   ABK51947 standard; cDNA; 2142 BP.
XX
AC   ABK51947;
XX
DT   27-AUG-2002  (first entry)
XX
DE   cDNA encoding human G-protein coupled receptor HGPRBMY5 splice variant.
XX
KW   Human; G-protein coupled receptor; GPCR; HGPRBMY5; colon; brain;
KW   ovary; thymus; lung; immune system; cancer; immune disorder;
KW   neurological disorder; infection; human immunodeficiency virus; HIV;
KW   antiallergic; antiasthmatic; dermatological; antiarteriosclerotic;
KW   antiulcer; antidiabetic; nephrotropic; osteopathic; antiarthritic;
KW   antiinflammatory; antirheumatic; antithyroid; cytostatic; vulnerary;
KW   virucide; antibacterial; antifungal; antiparasitic; protozoacide;
KW   antihelminthic; nootropic; neuroprotective; antidepressant;
KW   anticonvulsant; antiparkinsonian; neuroleptic; anti-HIV; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..2142
FT                   /*tag= a
FT                   /product= "GPCR HGPRBMY5"
XX
PN   WO200226824-A2.
XX
PD   04-APR-2002.
XX
PF   26-SEP-2001; 2001WO-US30365.
XX
PR   27-SEP-2000; 2000US-235713P.
PR   16-JAN-2001; 2001US-261781P.
PR   19-JUL-2001; 2001US-306605P.
PR   03-AUG-2001; 2001US-310436P.
XX
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI   Feder JN,  Mintier G,  Ramanathan CS,  Hawken DR;
XX
DR   WPI; 2002-435196/46.
DR   P-PSDB; AAU97159.
XX
PT   Novel G protein-coupled receptor, HGPRBMY5 polypeptide, useful for
PT   treatment of AIDS, allergies, asthma, atherosclerosis, ulcerative
PT   colitis, atopic dermatitis, diabetes mellitus, glomerulonephritis,
PT   osteoarthritis -
XX
PS   Claim 1; Fig 5; 148pp; English.
XX
```

```
CC   The present invention relates to the isolation of a novel human
CC   G-protein coupled receptor (GPCR) (HGPRBMY5), and the polynucleotide
CC   sequence encoding it. The HGPRBMY5 polypeptide and polynucleotide
CC   are useful for preventing, treating or ameliorating a disease,
CC   disorder or condition related to the colon, brain, ovaries, thymus,
CC   lungs or immune system. They are particularly useful for the
CC   treatment or prevention of cancers, immune disorders, neurological
CC   disorders, and diseases related to the brain, ovaries, thymus or
CC   lungs. The polynucleotide sequence is useful for diagnosing or
CC   determining susceptibility to infections such as bacterial, fungal,
CC   protozoan and viral infections, particularly infections caused by
CC   human immunodeficiency virus (HIV or HIV-2). The present sequence
CC   encodes a splice variant of human GPCR HGPRBMY5.
XX
SQ   Sequence 2142 BP; 613 A; 435 C; 408 G; 686 T; 0 other;

  Query Match             100.0%;  Score 321;  DB 24;  Length 2142;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-86;
  Matches  321;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1663 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1722

Qy     61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1723 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1782

Qy    121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1783 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1842

Qy    181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1843 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1902

Qy    241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1903 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 1962

Qy    301 AACTTTTTTAAGGACAAGTTG 321
          |||||||||||||||||||||
Db   1963 AACTTTTTTAAGGACAAGTTG 1983


RESULT 8
ABL40197
ID   ABL40197 standard; cDNA; 2190 BP.
XX
AC   ABL40197;
XX
DT   23-MAY-2002  (first entry)
XX
DE   Human G protein-coupled receptor TGR17-6 encoding cDNA SEQ ID NO:16.
XX
KW   Human; G protein-coupled receptor; TGR17-6; nootropic; antiinflammatory;
KW   vasotropic; immunomodulator; cytostatic; gene therapy; protein therapy;
KW   neurological; inflammatory; circulatory; degenerative; immune system;
KW   digestive disease; cancer; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..2190
FT                   /*tag= a
FT                   /partial
FT                   /product= "TGR17-6"
FT                   /note= "no stop codon given"
XX
PN   WO200204640-A1.
XX
PD   17-JAN-2002.
XX
PF   06-JUL-2001; 2001WO-JP05878.
XX
PR   07-JUL-2000; 2000JP-0211989.
PR   18-DEC-2000; 2000JP-0383794.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
PI   Moriya T,  Ito T,  Shintani Y,  Miyajima N;
XX
DR   WPI; 2002-179706/23.
DR   P-PSDB; ABB06255.
XX
PT   G-protein coupled receptor protein TGR17 of human origin and DNA
PT   encoding it for diagnosis and treatment of cancer and circulatory and
PT   other diseases associated with its expression -
XX
PS   Claim 7; Page 135-136; 145pp; Japanese.
XX
CC   The present invention describes a human guanine nucleotide binding
CC   protein (G protein) coupled receptor protein designated TGR17, which has
CC   5 variant forms (TGR17-1 to TGR17-5) of the receptor protein having
CC   additional residues at the N-terminal. The TGR17 proteins have nootropic,
CC   antiinflammatory, vasotropic, immunomodulator and cytostatic activities.
CC   The TGR17 polynucleotides and protein can be used in gene therapy and
CC   protein therapy. G-protein coupled receptor proteins are cell membrane
CC   proteins mediating the cellular response to a large variety of signalling
CC   molecules. The TGR17 polynucleotides and proteins can be used in the
CC   diagnosis, treatment and prevention of diseases including neurological,
CC   inflammatory, circulatory, degenerative, immune system and digestive
CC   diseases and cancer. The present sequence encodes human TGR17-6 from the
CC   present invention.
XX
SQ   Sequence 2190 BP; 626 A; 457 C; 409 G; 698 T; 0 other;

  Query Match             100.0%;  Score 321;  DB 24;  Length 2190;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-86;
  Matches  321;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1714 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1773

Qy     61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1774 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1833

Qy    121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1834 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1893

Qy    181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1894 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1953

Qy    241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1954 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2013

Qy    301 AACTTTTTTAAGGACAAGTTG 321
          |||||||||||||||||||||
Db   2014 AACTTTTTTAAGGACAAGTTG 2034


RESULT 9
ABK51944
ID   ABK51944 standard; cDNA; 2214 BP.
XX
AC   ABK51944;
XX
DT   27-AUG-2002  (first entry)
XX
DE   cDNA encoding human G-protein coupled receptor HGPRBMY5.
XX
```

```
KW   Human; G-protein coupled receptor; GPCR; HGPRBMY5; colon; brain;
KW   ovary; thymus; lung; immune system; cancer; immune disorder;
KW   neurological disorder; infection; human immunodeficiency virus; HIV;
KW   antiallergic; antiasthmatic; dermatological; antiarteriosclerotic;
KW   antiulcer; antidiabetic; nephrotropic; osteopathic; antiarthritic;
KW   antiinflammatory; antirheumatic; antithyroid; cytostatic; vulnerary;
KW   virucide; antibacterial; antifungal; antiparasitic; protozoacide;
KW   antihelminthic; nootropic; neuroprotective; antidepressant;
KW   anticonvulsant; antiparkinsonian; neuroleptic; anti-HIV; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..2214
FT                   /*tag= a
FT                   /product= "GPCR HGPRBMY5"
XX
PN   WO200226824-A2.
XX
PD   04-APR-2002.
XX
PF   26-SEP-2001; 2001WO-US30365.
XX
PR   27-SEP-2000; 2000US-235713P.
PR   16-JAN-2001; 2001US-261781P.
PR   19-JUL-2001; 2001US-306605P.
PR   03-AUG-2001; 2001US-310436P.
XX
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI   Feder JN,  Mintier G,  Ramanathan CS,  Hawken DR;
XX
DR   WPI; 2002-435196/46.
DR   P-PSDB; AAU97158.
XX
PT   Novel G protein-coupled receptor, HGPRBMY5 polypeptide, useful for
PT   treatment of AIDS, allergies, asthma, atherosclerosis, ulcerative
PT   colitis, atopic dermatitis, diabetes mellitus, glomerulonephritis,
PT   osteoarthritis -
XX
PS   Claim 1; Fig 1; 148pp; English.
XX
CC   The present invention relates to the isolation of a novel human
CC   G-protein coupled receptor (GPCR) (HGPRBMY5), and the polynucleotide
CC   sequence encoding it. The HGPRBMY5 polypeptide and polynucleotide
CC   are useful for preventing, treating or ameliorating a disease,
CC   disorder or condition related to the colon, brain, ovaries, thymus,
CC   lungs or immune system. They are particularly useful for the
CC   treatment or prevention of cancers, immune disorders, neurological
CC   disorders, and diseases related to the brain, ovaries, thymus or
CC   lungs. The polynucleotide sequence is useful for diagnosing or
CC   determining susceptibility to infections such as bacterial, fungal,
CC   protozoan and viral infections, particularly infections caused by
CC   human immunodeficiency virus (HIV or HIV-2). The present sequence
CC   encodes human GPCR HGPRBMY5.
XX
SQ   Sequence 2214 BP; 638 A; 457 C; 416 G; 703 T; 0 other;

  Query Match             100.0%;  Score 321;  DB 24;  Length 2214;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-86;
  Matches  321;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1735 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1794

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1795 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1854

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1855 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1914

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1915 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1974

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1975 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2034

Qy     301 AACTTTTTTAAGGACAAGTTG 321
          |||||||||||||||||||||
Db    2035 AACTTTTTTAAGGACAAGTTG 2055


RESULT 10
ABL40196
ID   ABL40196 standard; cDNA; 2262 BP.
XX
AC   ABL40196;
XX
DT   23-MAY-2002  (first entry)
XX
DE   Human G protein-coupled receptor TGR17-5 encoding cDNA SEQ ID NO:14.
XX
KW   Human; G protein-coupled receptor; TGR17-5; nootropic; antiinflammatory;
KW   vasotropic; immunomodulator; cytostatic; gene therapy; protein therapy;
KW   neurological; inflammatory; circulatory; degenerative; immune system;
KW   digestive disease; cancer; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..2262
FT                   /*tag= a
FT                   /partial
FT                   /product= "TGR17-5"
FT                   /note= "no stop codon given"
XX
PN   WO200204640-A1.
XX
PD   17-JAN-2002.
XX
PF   06-JUL-2001; 2001WO-JP05878.
XX
PR   07-JUL-2000; 2000JP-0211989.
PR   18-DEC-2000; 2000JP-0383794.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
PI   Moriya T,  Ito T,  Shintani Y,  Miyajima N;
XX
DR   WPI; 2002-179706/23.
DR   P-PSDB; ABB06254.
XX
PT   G-protein coupled receptor protein TGR17 of human origin and DNA
PT   encoding it for diagnosis and treatment of cancer and circulatory and
PT   other diseases associated with its expression -
XX
PS   Claim 7; Page 130-131; 145pp; Japanese.
XX
CC   The present invention describes a human guanine nucleotide binding
CC   protein (G protein) coupled receptor protein designated TGR17, which has
CC   5 variant forms (TGR17-1 to TGR17-5) of the receptor protein having
CC   additional residues at the N-terminal. The TGR17 proteins have nootropic,
CC   antiinflammatory, vasotropic, immunomodulator and cytostatic activities.
CC   The TGR17 polynucleotides and protein can be used in gene therapy and
CC   protein therapy. G-protein coupled receptor proteins are cell membrane
CC   proteins mediating the cellular response to a large variety of signalling
CC   molecules. The TGR17 polynucleotides and proteins can be used in the
CC   diagnosis, treatment and prevention of diseases including neurological,
CC   inflammatory, circulatory, degenerative, immune system and digestive
CC   diseases and cancer. The present sequence encodes human TGR17-5 from the
CC   present invention.
```

```
XX
SQ   Sequence 2262 BP; 650 A; 467 C; 421 G; 724 T; 0 other;

  Query Match             100.0%;  Score 321;  DB 24;  Length 2262;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-86;
  Matches  321;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1786 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1845

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1846 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1905

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1906 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1965

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1966 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 2025

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2026 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2085

Qy     301 AACTTTTTTAAGGACAAGTTG 321
          |||||||||||||||||||||
Db    2086 AACTTTTTTAAGGACAAGTTG 2106


RESULT 11
AAH51001/c
ID   AAH51001 standard; DNA; 1018 BP.
XX
AC   AAH51001;
XX
DT   28-AUG-2001  (first entry)
XX
DE   Human nGPCR57 coding sequence.
XX
KW   G protein-coupled receptor; nGPCR; seven transmembrane receptor;
KW   signal transduction; schizophrenia; thyroid disorder; renal failure;
KW   rheumatoid arthritis; CNS disorder; infection; metabolic disease;
KW   cardiovascular disease; proliferative disorder; hormonal disorder;
KW   neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
KW   attention deficit-hyperactivity disorder/attention deficit disorder;
KW   Parkinson's disease; migraine; senile dementia; inflammatory disease;
KW   rheumatoid arthritis; autoimmune disorder; respiratory ailment;
KW   neuroprotective; ds.
XX
OS   Homo sapiens.
XX
PN   WO200136473-A2.
XX
PD   25-MAY-2001.
XX
PF   16-NOV-2000; 2000WO-US31581.
XX
PR   16-NOV-1999;   99US-0165838.
PR   17-NOV-1999;   99US-0166071.
PR   19-NOV-1999;   99US-0166678.
PR   28-DEC-1999;   99US-0173396.
PR   22-FEB-2000; 2000US-0184129.
PR   28-FEB-2000; 2000US-0185421.
PR   28-FEB-2000; 2000US-0185554.
PR   02-MAR-2000; 2000US-0186530.
PR   03-MAR-2000; 2000US-0186811.
PR   09-MAR-2000; 2000US-0188114.
PR   17-MAR-2000; 2000US-0190310.
PR   21-MAR-2000; 2000US-0190800.
PR   20-APR-2000; 2000US-0198568.
PR   02-MAY-2000; 2000US-0201190.
PR   08-MAY-2000; 2000US-0203111.
PR   25-MAY-2000; 2000US-0207094.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Vogeli G,  Wood LS,  Parodi LA,  Hiebsch RR,  Lind P,  Slightom J;
PI   Schellin KA,  Kaytes PS,  Bannigan CM,  Ruff V,  Sejlitz T,  Huff RM;
XX
DR   WPI; 2001-389826/41.
DR   P-PSDB; AAG80961.
XX
PT   New G protein-coupled receptor (nGPCR-x) and its encoding
PT   polynucleotide useful for diagnosing and treating e.g. schizophrenia -
XX
PS   Claim 4; Page 86; 261pp; English.
XX
CC   The present invention relates to novel G protein-coupled receptors
CC   (nGPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
CC   28, 31-38, 40, 41, 53-60) and their coding sequences. The present
CC   sequence is the coding sequence for one such G protein-coupled receptor.
CC   GPCRs are also known as seven transmembrane receptors and function in
CC   signal transduction. The nGPCRx coding sequences are useful for
CC   screening a human to diagnose a disorder affecting the brain or a genetic
CC   predisposition, specifically schizophrenia. nGPCRx are useful for
CC   identifying compounds useful for treating schizophrenia. Detection of
CC   nGPCRx in a sample is useful as a diagnostic tool for diseases or
CC   disorders e.g. thyroid disorders, renal failure, rheumatoid arthritis,
CC   CNS disorders, infections such as HIV-1, metabolic and cardiovascular
CC   diseases, proliferative disorders and hormonal disorders. Modulators of
CC   nGPCRx activity have the utility for treating neurological disorders,
CC   including schizophrenia, ADHD/ADD (attention deficit-hyperactivity
CC   disorder/attention deficit disorder), and neuronal disorders such as
CC   Alzheimer's disease, Parkinson's disease, migraine and senile dementia.
CC   Additional disorders include inflammatory conditions (e.g. Crohn's
CC   disease), rheumatoid arthritis, autoimmune disorders, cancers,
CC   respiratory ailments such as asthma, and inflammatory diseases e.g.
CC   inflammatory bowel disease.
XX
SQ   Sequence 1018 BP; 332 A; 179 C; 189 G; 318 T; 0 other;

  Query Match             68.5%;  Score 220;  DB 22;  Length 1018;
  Best Local Similarity   100.0%;  Pred. No. 4.1e-56;
  Matches  220;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     539 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 480

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     479 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 420

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     419 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 360

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAG 220
          |||||||||||||||||||||||||||||||||||||||
Db     359 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAG 320


RESULT 12
AAI99584
ID   AAI99584 standard; cDNA; 530 BP.
XX
AC   AAI99584;
XX
DT   04-JAN-2002  (first entry)
XX
DE   Human expressed polynucleotide SEQ ID NO 47.
XX
KW   Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
```

KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine;
KW ss.
XX
OS Homo sapiens.
XX
PN WO200155387-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01310.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.

```
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-465573/50.
DR   P-PSDB; AAM99972.
XX
PT   Isolated digestive system associated polypeptide for treating,
PT   preventing and/ or prognosing disorders related to the digestive system
PT   including digestive system cancers and also for testing and detection
PT   e.g. diagnosis -
XX
PS   Claim 1; SEQ ID NO 47; 509pp + Sequence Listing; English.
XX
CC   The invention relates to novel genes (AAI99548-AAI99604) and proteins
CC   (AAM99936-AAM99984) useful for preventing, treating or ameliorating
CC   medical conditions e.g. by protein or gene therapy. The genes are
CC   isolated from a range of human tissues disclosed in the specification.
CC   The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC   in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC   and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC   marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC   (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC   haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC   disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC   colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC   (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC   epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC   and parasitic infections.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other;

  Query Match             51.8%;  Score 166.4;  DB 22;  Length 530;
  Best Local Similarity   70.0%;  Pred. No. 4.2e-40;
  Matches  224;  Conservative    0;  Mismatches   96;  Indels    0;  Gaps    0;

Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
           ||| | || |||   || ||| |||||||| || |||||||||   |  ||||| | |
Db      32 GGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGT 91

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
            ||||  |||  ||| |     | || ||| || |||||    ||  || |||| ||
Db      92 GTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATC 151

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
            |||| || ||||| || |||||||| ||  |||||||  | ||||||||  ||  | |||
Db     152 CTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTT 211

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
           ||||| ||| | ||||| ||  | | ||| ||||||||||  || || || || ||| ||
Db     212 GTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTA 271

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
           ||||||||| | || ||  |||||||||||||||| ||||| |||||||||| || ||
Db     272 GTGATTTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACA 331

Qy     301 AACTTTTTTAAGGACAAGTT 320
           |     ||||| || | | |
Db     332 AGACCATTTAAAGAAATGAT 351


RESULT 13
AAL35644
ID   AAL35644 standard; cDNA; 530 BP.
XX
AC   AAL35644;
XX
DT   08-JAN-2002  (first entry)
XX
DE   Human musculoskeletal system related polynucleotide SEQ ID NO 986.
XX
KW   Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW   antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW   vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW   cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; human; secreted protein;
KW   musculoskeletal system; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155367-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01338.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
```

PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-451937/48.
DR P-PSDB; ABB04062.
XX
PT Isolated polypeptide for treating, preventing and/ or prognosing
PT disorders related to the musculoskeletal system including
PT musculoskeletal cancers and also for testing and detection e.g.
PT diagnosis -
XX
PS Claim 1; SEQ ID NO 986; 781pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (AAL34669-AAL37666) and proteins
CC (ABB03087-ABB04109) associated with the musculoskeletal system useful
CC for preventing, treating or ameliorating medical conditions e.g. by
CC protein or gene therapy. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids, proteins,
CC antibodies and (ant)agonists are useful in the diagnosis, treatment
CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and
CC other cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC (c) cardiovascular disorders such as myocardial ischaemias; (d) wound
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC and (f) infectious diseases such as viral, bacterial, fungal and
CC parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other;

Query Match 51.8%; Score 166.4; DB 22; Length 530;
Best Local Similarity 70.0%; Pred. No. 4.2e-40;
Matches 224; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

```
Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
           ||| | || |||   || ||| ||||||| || |||||||||   |  ||||| | |
Db      32 GGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGT 91

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
             ||||  |||  ||| |     | || ||| || |||||    ||  || |||| ||
Db      92 GTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATC 151

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
             |||| || ||||| || |||||||| ||  |||||||  | ||||||||| ||  | |||
Db     152 CTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTT 211

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
           ||||| ||| | ||||| ||  | | ||| ||||||||||  || || || || ||| ||
Db     212 GTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTA 271

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
           ||||||||| | || ||  |||||||||||||||| ||||| ||||||||||| || ||
Db     272 GTGATTTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACA 331
```

```
Qy     301 AACTTTTTTAAGGACAAGTT 320
           |     ||||| || | | |
Db     332 AGACCATTTAAAGAAATGAT 351


RESULT 14
ABA06471
ID   ABA06471 standard; cDNA; 530 BP.
XX
AC   ABA06471;
XX
DT   10-JAN-2002  (first entry)
XX
DE   Human cDNA SEQ ID NO: 137.
XX
KW   Human; gene therapy; neural disorder; immune system disorder;
KW   muscular disorder; reproductive disorder; gastrointestinal disorder;
KW   pulmonary disorder; cardiovascular disorder; renal disorder;
KW   proliferative disorder; inflammation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200154474-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01349.
XX
PR   31-JAN-2000; 2000US-179065P.
PR   04-FEB-2000; 2000US-180628P.
PR   24-FEB-2000; 2000US-184664P.
PR   02-MAR-2000; 2000US-186350P.
PR   16-MAR-2000; 2000US-189874P.
PR   17-MAR-2000; 2000US-190076P.
PR   18-APR-2000; 2000US-198123P.
PR   19-MAY-2000; 2000US-205515P.
PR   07-JUN-2000; 2000US-209467P.
PR   28-JUN-2000; 2000US-214886P.
PR   30-JUN-2000; 2000US-215135P.
PR   07-JUL-2000; 2000US-216647P.
PR   07-JUL-2000; 2000US-216880P.
PR   11-JUL-2000; 2000US-217487P.
PR   11-JUL-2000; 2000US-217496P.
PR   14-JUL-2000; 2000US-218290P.
PR   26-JUL-2000; 2000US-220963P.
PR   26-JUL-2000; 2000US-220964P.
PR   14-AUG-2000; 2000US-224518P.
PR   14-AUG-2000; 2000US-224519P.
PR   14-AUG-2000; 2000US-225213P.
PR   14-AUG-2000; 2000US-225214P.
PR   14-AUG-2000; 2000US-225266P.
PR   14-AUG-2000; 2000US-225267P.
PR   14-AUG-2000; 2000US-225268P.
PR   14-AUG-2000; 2000US-225270P.
PR   14-AUG-2000; 2000US-225447P.
PR   14-AUG-2000; 2000US-225757P.
PR   14-AUG-2000; 2000US-225758P.
PR   14-AUG-2000; 2000US-225759P.
PR   18-AUG-2000; 2000US-226279P.
PR   22-AUG-2000; 2000US-226681P.
PR   22-AUG-2000; 2000US-226868P.
PR   22-AUG-2000; 2000US-227182P.
PR   23-AUG-2000; 2000US-227009P.
PR   30-AUG-2000; 2000US-228924P.
PR   01-SEP-2000; 2000US-229287P.
PR   01-SEP-2000; 2000US-229343P.
PR   01-SEP-2000; 2000US-229344P.
PR   01-SEP-2000; 2000US-229345P.
PR   05-SEP-2000; 2000US-229509P.
PR   05-SEP-2000; 2000US-229513P.
PR   06-SEP-2000; 2000US-230437P.
PR   06-SEP-2000; 2000US-230438P.
PR   08-SEP-2000; 2000US-231242P.
PR   08-SEP-2000; 2000US-231243P.
PR   08-SEP-2000; 2000US-231244P.
PR   08-SEP-2000; 2000US-231413P.
PR   08-SEP-2000; 2000US-231414P.
PR   08-SEP-2000; 2000US-232080P.
PR   08-SEP-2000; 2000US-232081P.
PR   12-SEP-2000; 2000US-231968P.
PR   14-SEP-2000; 2000US-232397P.
PR   14-SEP-2000; 2000US-232398P.
PR   14-SEP-2000; 2000US-232399P.
PR   14-SEP-2000; 2000US-232400P.
PR   14-SEP-2000; 2000US-232401P.
PR   14-SEP-2000; 2000US-233063P.
PR   14-SEP-2000; 2000US-233064P.
PR   14-SEP-2000; 2000US-233065P.
PR   21-SEP-2000; 2000US-234223P.
PR   21-SEP-2000; 2000US-234274P.
PR   25-SEP-2000; 2000US-234997P.
PR   25-SEP-2000; 2000US-234998P.
PR   26-SEP-2000; 2000US-235484P.
PR   27-SEP-2000; 2000US-235834P.
PR   27-SEP-2000; 2000US-235836P.
PR   29-SEP-2000; 2000US-236327P.
PR   29-SEP-2000; 2000US-236367P.
PR   29-SEP-2000; 2000US-236368P.
PR   29-SEP-2000; 2000US-236369P.
PR   29-SEP-2000; 2000US-236370P.
PR   02-OCT-2000; 2000US-236802P.
PR   02-OCT-2000; 2000US-237037P.
PR   02-OCT-2000; 2000US-237038P.
PR   02-OCT-2000; 2000US-237039P.
PR   02-OCT-2000; 2000US-237040P.
PR   13-OCT-2000; 2000US-239935P.
PR   13-OCT-2000; 2000US-239937P.
PR   20-OCT-2000; 2000US-240960P.
PR   20-OCT-2000; 2000US-241221P.
PR   20-OCT-2000; 2000US-241785P.
PR   20-OCT-2000; 2000US-241786P.
PR   20-OCT-2000; 2000US-241787P.
PR   20-OCT-2000; 2000US-241808P.
PR   20-OCT-2000; 2000US-241809P.
PR   20-OCT-2000; 2000US-241826P.
PR   01-NOV-2000; 2000US-244617P.
PR   08-NOV-2000; 2000US-246474P.
PR   08-NOV-2000; 2000US-246475P.
PR   08-NOV-2000; 2000US-246476P.
PR   08-NOV-2000; 2000US-246477P.
PR   08-NOV-2000; 2000US-246478P.
PR   08-NOV-2000; 2000US-246523P.
PR   08-NOV-2000; 2000US-246524P.
PR   08-NOV-2000; 2000US-246525P.
PR   08-NOV-2000; 2000US-246526P.
PR   08-NOV-2000; 2000US-246527P.
PR   08-NOV-2000; 2000US-246528P.
PR   08-NOV-2000; 2000US-246532P.
PR   08-NOV-2000; 2000US-246609P.
PR   08-NOV-2000; 2000US-246610P.
PR   08-NOV-2000; 2000US-246611P.
PR   08-NOV-2000; 2000US-246613P.
PR   17-NOV-2000; 2000US-249207P.
PR   17-NOV-2000; 2000US-249208P.
PR   17-NOV-2000; 2000US-249209P.
PR   17-NOV-2000; 2000US-249210P.
PR   17-NOV-2000; 2000US-249211P.
PR   17-NOV-2000; 2000US-249212P.
PR   17-NOV-2000; 2000US-249213P.
PR   17-NOV-2000; 2000US-249214P.
PR   17-NOV-2000; 2000US-249215P.
PR   17-NOV-2000; 2000US-249216P.
PR   17-NOV-2000; 2000US-249217P.
PR   17-NOV-2000; 2000US-249218P.
PR   17-NOV-2000; 2000US-249244P.
```

```
PR   17-NOV-2000; 2000US-249245P.
PR   17-NOV-2000; 2000US-249264P.
PR   17-NOV-2000; 2000US-249265P.
PR   17-NOV-2000; 2000US-249297P.
PR   17-NOV-2000; 2000US-249299P.
PR   17-NOV-2000; 2000US-249300P.
PR   01-DEC-2000; 2000US-250160P.
PR   01-DEC-2000; 2000US-250391P.
PR   05-DEC-2000; 2000US-251030P.
PR   05-DEC-2000; 2000US-251988P.
PR   05-DEC-2000; 2000US-256719P.
PR   06-DEC-2000; 2000US-251479P.
PR   08-DEC-2000; 2000US-251856P.
PR   08-DEC-2000; 2000US-251868P.
PR   08-DEC-2000; 2000US-251869P.
PR   08-DEC-2000; 2000US-251989P.
PR   08-DEC-2000; 2000US-251990P.
PR   11-DEC-2000; 2000US-254097P.
PR   05-JAN-2001; 2001US-259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-476161/51.
DR   P-PSDB; ABB10249.
XX
PT   Isolated nucleic acid molecule encoding an inflammation-associated
PT   polypeptide is used in preventing, treating or ameliorating a medical
PT   condition  -
XX
PS   Claim 1; SEQ ID NO: 137; 859pp + Sequence Listing; English.
XX
CC   The present invention provides human cDNAs, proteins and related genomic
CC   DNAs. These can be used in the treatment of neural, immune system,
CC   muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC   renal and proliferative disorders and inflammation. The present sequence
CC   is a cDNA of the invention.
XX
SQ   Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other;

  Query Match             51.8%;  Score 166.4;  DB 22;  Length 530;
  Best Local Similarity   70.0%;  Pred. No. 4.2e-40;
  Matches  224;  Conservative    0;  Mismatches   96;  Indels    0;  Gaps    0;

Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
           ||| | || |||   || ||| ||||||| || |||||||||   |  ||||| | |
Db      32 GGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGT 91

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
             ||||  |||  ||| |     | || ||| || |||||    ||  || |||| ||
Db      92 GTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATC 151

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
             |||| || ||||| || |||||||| ||  |||||||  | ||||||||  ||  | |||
Db     152 CTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTT 211

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
           ||||| ||| | ||||| ||  | | ||| ||||||||||  || || || || ||| ||
Db     212 GTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTA 271

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
           |||||||||  | || ||  |||||||||||||||| ||||| |||||||||| || ||
Db     272 GTGATTTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACA 331

Qy     301 AACTTTTTTAAGGACAAGTT 320
           |     ||||| || | | |
Db     332 AGACCATTTAAAGAAATGAT 351


RESULT 15
AAS28950
ID   AAS28950 standard; cDNA; 530 BP.
XX
AC   AAS28950;
XX
DT   21-NOV-2001  (first entry)
XX
DE   cDNA encoding for human uterine motility-association polypeptide #15.
XX
KW   Human; uterine motility-association disorder; uterus; pregnancy;
KW   labour; menstrual cycle; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155201-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01317.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
```

PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-488777/53.
DR P-PSDB; AAU18108.
XX
PT Isolated polypeptide and nucleic acid molecules for treating,
PT preventing and/or prognosing disorders related to uterine motility
PT e.g. disorders associated with pregnancy and the menstrual cycle -
XX
PS Claim 4; SEQ ID No 25; 524pp; English.
XX
CC The present invention relates to the isolation of novel human
CC uterine motility-association polypeptides (AAU18094-AAU18152),
CC and cDNA and genomic sequences encoding for these polypeptides.
CC The sequences of the invention are useful in the diagnosis,
CC treatment, prevention and/or prognosis of diseases associated
CC with uterine motility such as pregnancy and labour, and menstrual
CC disorders. The polynucleotide sequences of the invention are also
CC useful in gene therapy. AAS28936-AAS28994 represent cDNA sequences
CC encoding for novel human uterine motility-association polypeptides.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other;

Query Match 51.8%; Score 166.4; DB 22; Length 530;
Best Local Similarity 70.0%; Pred. No. 4.2e-40;
Matches 224; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

```
Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
           ||| | || |||   || ||| ||||||| || |||||||||.  |  ||||| | |
Db      32 GGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGT 91

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
             ||||  |||  ||| |     | || ||| || |||||    ||  || |||| ||
Db      92 GTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATC 151

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
             |||| || ||||| || |||||||| ||  |||||||  | |||||||| ||  | |||
Db     152 CTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTT 211

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
           ||||| ||| | |||||  ||  | | ||| |||||||||||  || || || || ||| ||
Db     212 GTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTA 271

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
           |||||||||  |  || ||  |||||||||||||||| ||||| |||||||||| || ||
Db     272 GTGATTTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACA 331

Qy     301 AACTTTTTTAAGGACAAGTT 320
           |     ||||| || | | |
Db     332 AGACCATTTAAAGAAATGAT 351
```

Search completed: May 26, 2003, 19:02:34
Job time : 244 secs

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 18:47:57 ; Search time 52 Seconds
(without alignments)
1893.139 Million cell updates/sec

Title: US-09-930-312-1
Perfect score: 321
Sequence: 1 ggtgtgaacttgctggcttt.....actttttaaggacaagttg 321

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	40.8	12.7	7218	1	US-08-232-463-14	Sequence 14, Appl
	2	34.8	10.8	1831	1	US-08-101-435-3	Sequence 3, Appli
	3	32.4	10.1	957	1	US-08-349-696-26	Sequence 26, Appl
	4	32.4	10.1	957	1	US-08-233-009-26	Sequence 26, Appl
	5	32.4	10.1	957	1	US-08-560-231-26	Sequence 26, Appl
	6	32.4	10.1	957	4	US-09-080-704A-26	Sequence 26, Appl
	7	30.8	9.6	1695	4	US-09-149-476-299	Sequence 299, App
c	8	30.6	9.5	51952	3	US-08-947-823-1	Sequence 1, Appli
	9	30.2	9.4	7486	3	US-08-475-886-5	Sequence 5, Appli
	10	30.2	9.4	7486	4	US-08-397-232-3	Sequence 3, Appli
	11	30.2	9.4	7486	4	US-09-653-499-5	Sequence 5, Appli
	12	30.2	9.4	7488	3	US-08-475-886-3	Sequence 3, Appli
	13	30.2	9.4	7488	4	US-09-653-499-3	Sequence 3, Appli
	14	30	9.3	1655	2	US-08-945-296-2	Sequence 2, Appli
	15	30	9.3	1655	3	US-09-405-112-2	Sequence 2, Appli
c	16	30	9.3	7521	4	US-09-004-838-116	Sequence 116, App
c	17	29.8	9.3	1400	4	US-09-117-257-10	Sequence 10, Appl
c	18	29.8	9.3	1400	4	US-08-945-476-10	Sequence 10, Appl
c	19	29.8	9.3	1400	4	US-09-489-352-10	Sequence 10, Appl
c	20	29.8	9.3	2653	2	US-08-589-711-1	Sequence 1, Appli
c	21	29.8	9.3	2653	4	US-09-221-938-1	Sequence 1, Appli
c	22	29.8	9.3	2653	4	US-08-945-476-7	Sequence 7, Appli
c	23	29.8	9.3	2656	4	US-09-117-257-7	Sequence 7, Appli
c	24	29.8	9.3	2656	4	US-09-489-352-7	Sequence 7, Appli
	25	29.6	9.2	981	1	US-08-349-696-20	Sequence 20, Appl
	26	29.6	9.2	981	1	US-08-233-009-20	Sequence 20, Appl
	27	29.6	9.2	981	1	US-08-560-231-20	Sequence 20, Appl
	28	29.6	9.2	981	4	US-09-080-704A-20	Sequence 20, Appl
	29	29.6	9.2	2180	1	US-07-918-314-3	Sequence 3, Appli
	30	29.6	9.2	2900	1	US-07-918-314-5	Sequence 5, Appli
	31	29.4	9.2	1302	2	US-08-529-600D-3	Sequence 3, Appli
	32	29.4	9.2	1302	2	US-08-973-275-6	Sequence 6, Appli
	33	29.4	9.2	1302	3	US-09-122-632-3	Sequence 3, Appli
c	34	29.4	9.2	4084	2	US-08-568-459A-1	Sequence 1, Appli
c	35	29.4	9.2	4084	2	US-08-487-826B-1	Sequence 1, Appli
c	36	29.4	9.2	4084	4	US-09-210-288-1	Sequence 1, Appli
c	37	29.4	9.2	4084	6	5198347-5	Patent No. 5198347
c	38	29.4	9.2	9909	4	US-08-961-527-12	Sequence 12, Appl
c	39	29.2	9.1	2047	4	US-09-453-702B-222	Sequence 222, App
	40	29	9.0	883	1	US-07-828-798C-4	Sequence 4, Appli
	41	29	9.0	883	2	US-08-315-868A-4	Sequence 4, Appli
	42	29	9.0	883	3	US-08-495-819B-4	Sequence 4, Appli
c	43	29	9.0	1882	1	US-08-458-120-1	Sequence 1, Appli
c	44	29	9.0	1882	2	US-08-867-970-1	Sequence 1, Appli
c	45	29	9.0	1882	4	US-09-326-217-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match 12.7%; Score 40.8; DB 1; Length 7218;

```
  Best Local Similarity   2.9%;  Pred. No. 0.014;
  Matches    9;  Conservative  176;  Mismatches  123;  Indels    0;  Gaps    0;

Qy       2 GTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCCA 61
           | | || |||||   :::::::: :: :: : :::::::: : :: :: : :::: :::: 
Db    1052 GAGGGAGCTTGCGATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1111

Qy      62 TTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCTG 121
           :::       :: :::: :   :: :       :       :: ::::          :  :: 
Db    1112 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1171

Qy     122 TTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTTG 181
           :: :   :: :::::::::: :  : :::::  : :: ::: ::   ::::: : :::  
Db    1172 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1231

Qy     182 TAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATAG 241
           :  ::    ::::::::::::::::   :       : ::   : :  :  ::::::    :  
Db    1232 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1291

Qy     242 TGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACCA 301
           :  :::::::::::::  ::   :  : ::::   :::  :::::: : ::::: :  ::  
Db    1292 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1351

Qy     302 ACTTTTTT 309
           :::::::
Db    1352 YYYYYYYY 1359


RESULT 2
US-08-101-435-3
; Sequence 3, Application US/08101435
; Patent No. 5441883
;  GENERAL INFORMATION:
;    APPLICANT:  Civelli, Olivier
;    APPLICANT:  Zhou, Qun-yong
;    TITLE OF INVENTION:  A No. 5441883el Adenosine Receptor and Uses
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Allegretti & Witcoff, Ltd.
;      STREET:  10 South Wacker Drive, Suite 3000
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/101,435
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/847,563
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McDonnell, John J
;      REGISTRATION NUMBER:  26,949
;      REFERENCE/DOCKET NUMBER:  91,708
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-715-1000
;      TELEFAX:  312-715-1234
;      TELEX:  910-221-5317
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1831 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  316..1276
US-08-101-435-3

  Query Match            10.8%;  Score 34.8;  DB 1;  Length 1831;
  Best Local Similarity   50.0%;  Pred. No. 0.5;
  Matches   87;  Conservative    0;  Mismatches   87;  Indels    0;  Gaps    0;

Qy      84 CACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTTAT 143
           || |||      |    | ||  ||  | ||| |       || ||     | ||| |  |
Db     966 CAGAGAGACGCGTGCATTTTACGGTCGGGAGTTCAAGACCGCTAAGTCCCTGTTTCTGGT 1025

Qy     144 AGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCTT 203
             | ||||    |||| | |||||| | ||| | |    |  | ||  |  |||||  |||
Db    1026 TCTCTTCTTGTTTGCCTTGTGCTGGCTGCCTTTGTCCATCATCAATTTTGTTTCCTACTT 1085

Qy     204 CCGGGTGGAAATACCAGACACAATGACTTCCTGGATAGTGATTTTTTTCCTTCC 257
             ||| | |||||||| |     |  | |  |    |  | || | || | |
Db    1086 TAATGTGAAGATACCAGAGATTGCAATGTGCCTGGGCATCCTGTTGTCCCATGC 1139


RESULT 3
US-08-349-696-26
; Sequence 26, Application US/08349696
; Patent No. 5599671
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobson, Marlene A
;    APPLICANT:  Johnson, Robert G
;    APPLICANT:  Luneau, Christopher J
;    APPLICANT:  Salvatore, Christopher A
;    TITLE OF INVENTION:  Human Adenosine Receptors
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  P.O. Box 2000
;      CITY:  Rahway
;      STATE:  NJ
;      COUNTRY:  United States
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Macintosh IIci
;      OPERATING SYSTEM:  Macintosh
;      SOFTWARE:  Microsoft Word 5.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/349,696
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  us/08/005945
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meredith, Roy D.
;      REGISTRATION NUMBER:  30,777
;      REFERENCE/DOCKET NUMBER:  18699IA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (908)594-4678
;      TELEFAX:  (908)594-4720
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  957 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-349-696-26

  Query Match            10.1%;  Score 32.4;  DB 1;  Length 957;
  Best Local Similarity   49.1%;  Pred. No. 2;
  Matches  115;  Conservative    0;  Mismatches  116;  Indels    3;  Gaps    1;

Qy      83 CCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTTA 142
           ||| |||   | |    | || |||| | ||| |        || ||    || ||| |
Db     644 CCAAAGAGACAGGTGCATTTTATGGACGGGAGTTCAAGACGGCTAAGTCCTTGTTTCTGG 703
```

```
Qy    143 TAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCT 202
          |  | ||||    |||  | |  ||| | ||| ||| | |  | ||   | |   |  ||
Db    704 TTCTTTTCTTGTTTGCTCTGTCATGGCTGCCTTTATCTATCATCAACTGCATCATCTACT 763

Qy    203 TCCGGGTGGAAATACCAGACACAATGACTTCCTGGATAGTGATTTTTTTCCTTCCAGTTA 262
          |    | ||  ||||| |     ||  | |   || | ||  |  |    |  |  |
Db    764 TTAATGGTGAGGTACCACAGCTTGTGCTGTAC---ATGGGCATCCTGCTGTCCCATGCCA 820

Qy    263 ACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACCAACTTTTTTAAGGACA 316
          ||       |||| || ||| ||||| |    | ||  ||    || ||||| |
Db    821 ACTCCATGATGAACCCTATCGTCTATGCCTATAAAATAAAGAAGTTCAAGGAAA 874


RESULT 4
US-08-233-009-26
; Sequence 26, Application US/08233009
; Patent No. 5646156
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobson, Marlene A
;    APPLICANT:  Johnson, Robert G
;    APPLICANT:  Salvatore, Christopher A
;    TITLE OF INVENTION:  INHIBITION OF EOSINOPHIL
;    TITLE OF INVENTION:  ACTIVATION THROUGH A3 ADENOSINE RECEPTOR ANTAGONISM
;    NUMBER OF SEQUENCES:  56
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  P.O.Box 2000
;      CITY:  Rahway
;      STATE:  New Jersey
;      COUNTRY:  United States
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/233,009
;      FILING DATE:  25-APR-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bencen, Gerard H
;      REGISTRATION NUMBER:  35,746
;      REFERENCE/DOCKET NUMBER:  19219
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (908) 594-3901
;      TELEFAX:  (908)594-4720
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  957 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-233-009-26

  Query Match            10.1%;  Score 32.4;  DB 1;  Length 957;
  Best Local Similarity   49.1%;  Pred. No. 2;
  Matches  115;  Conservative    0;  Mismatches  116;  Indels    3;  Gaps    1;

Qy     83 CCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTTA 142
          ||| |||   | |    | || |||| | ||| |       || ||    || ||| |
Db    644 CCAAAGAGACAGGTGCATTTTATGGACGGGAGTTCAAGACGGCTAAGTCCTTGTTTCTGG 703

Qy    143 TAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCT 202
          |  | ||||    |||  | |  ||| | ||| ||| | |  | ||   | |   |  ||
Db    704 TTCTTTTCTTGTTTGCTCTGTCATGGCTGCCTTTATCTATCATCAACTGCATCATCTACT 763

Qy    203 TCCGGGTGGAAATACCAGACACAATGACTTCCTGGATAGTGATTTTTTTCCTTCCAGTTA 262
          |    | ||  ||||| |     ||  | |   || | ||  |  |    |  |  |
Db    764 TTAATGGTGAGGTACCACAGCTTGTGCTGTAC---ATGGGCATCCTGCTGTCCCATGCCA 820

Qy    263 ACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACCAACTTTTTTAAGGACA 316
          ||       |||| || ||| ||||| |    | ||  ||    || ||||| |
Db    821 ACTCCATGATGAACCCTATCGTCTATGCCTATAAAATAAAGAAGTTCAAGGAAA 874


RESULT 5
US-08-560-231-26
; Sequence 26, Application US/08560231
; Patent No. 5817760
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobson, Marlene A
;    APPLICANT:  Johnson, Robert G
;    APPLICANT:  Luneau, Christopher J
;    APPLICANT:  Salvatore, Christopher A
;    TITLE OF INVENTION:  Human Adenosine Receptors
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  P.O. Box 2000
;      CITY:  Rahway
;      STATE:  NJ
;      COUNTRY:  United States
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Macintosh IIci
;      OPERATING SYSTEM:  Macintosh
;      SOFTWARE:  Microsoft Word 5.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/560,231
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meredith, Roy D.
;      REGISTRATION NUMBER:  30,777
;      REFERENCE/DOCKET NUMBER:  18699IA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (908)594-4678
;      TELEFAX:  (908)594-4720
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  957 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-560-231-26

  Query Match            10.1%;  Score 32.4;  DB 1;  Length 957;
  Best Local Similarity   49.1%;  Pred. No. 2;
  Matches  115;  Conservative    0;  Mismatches  116;  Indels    3;  Gaps    1;

Qy     83 CCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTTA 142
          ||| |||   | |    | || |||| | ||| |       || ||    || ||| |
Db    644 CCAAAGAGACAGGTGCATTTTATGGACGGGAGTTCAAGACGGCTAAGTCCTTGTTTCTGG 703

Qy    143 TAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCT 202
          |  | ||||    |||  | |  ||| | ||| ||| | |  | ||   | |   |  ||
Db    704 TTCTTTTCTTGTTTGCTCTGTCATGGCTGCCTTTATCTATCATCAACTGCATCATCTACT 763

Qy    203 TCCGGGTGGAAATACCAGACACAATGACTTCCTGGATAGTGATTTTTTTCCTTCCAGTTA 262
          |    | ||  ||||| |     ||  | |   || | ||  |  |    |  |  |
Db    764 TTAATGGTGAGGTACCACAGCTTGTGCTGTAC---ATGGGCATCCTGCTGTCCCATGCCA 820

Qy    263 ACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACCAACTTTTTTAAGGACA 316
          ||       |||| || ||| ||||| |    | ||  ||    || ||||| |
Db    821 ACTCCATGATGAACCCTATCGTCTATGCCTATAAAATAAAGAAGTTCAAGGAAA 874


RESULT 6
US-09-080-704A-26
```

```
; Sequence 26, Application US/09080704A
; Patent No. 6166181
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobson, Marlene A
;    APPLICANT:  Johnson, Robert G
;    APPLICANT:  Luneau, Christopher J
;    APPLICANT:  Salvatore, Christopher A
;    TITLE OF INVENTION:  Human Adenosine Receptors
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  P.O. Box 2000
;      CITY:  Rahway
;      STATE:  NJ
;      COUNTRY:  United States
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  Windows NT
;      SOFTWARE:  Word 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/080,704A
;      FILING DATE:  18 May 1998
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parr, Richard S.
;      REGISTRATION NUMBER:  32,586
;      REFERENCE/DOCKET NUMBER:  18699DB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (732)594-4958
;      TELEFAX:  (732)594-4720
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  957 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-080-704A-26

  Query Match             10.1%;  Score 32.4;  DB 4;  Length 957;
  Best Local Similarity   49.1%;  Pred. No. 2;
  Matches  115;  Conservative    0;  Mismatches  116;  Indels    3;  Gaps    1;

Qy      83 CCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTTA 142
           ||| |||   | |    | || |||| | ||| |       || ||    || ||| |
Db     644 CCAAAGAGACAGGTGCATTTTATGGACGGGAGTTCAAGACGGCTAAGTCCTTGTTTCTGG 703

Qy     143 TAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCT 202
           |  | ||||    |||  | |  ||| | ||| ||| | |  | ||   | |   |  ||
Db     704 TTCTTTTCTTGTTTGCTCTGTCATGGCTGCCTTTATCTATCATCAACTGCATCATCTACT 763

Qy     203 TCCGGGTGGAAATACCAGACACAATGACTTCCTGGATAGTGATTTTTTTCCTTCCAGTTA 262
           |    |  ||  ||||| |     ||   | |   || |  ||  |  |    |  | |
Db     764 TTAATGGTGAGGTACCACAGCTTGTGCTGTAC---ATGGGCATCCTGCTGTCCCATGCCA 820

Qy     263 ACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACCAACTTTTTTAAGGACA 316
           ||       |||| || ||| ||||| |    | ||  ||    || |||||
Db     821 ACTCCATGATGAACCCTATCGTCTATGCCTATAAAATAAAGAAGTTCAAGGAAA 874


RESULT 7
US-09-149-476-299
; Sequence 299, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 186 Human Secreted proteins
;  FILE REFERENCE: PZ002P1
;  CURRENT APPLICATION NUMBER: US/09/149,476
;  CURRENT FILING DATE: 1998-09-08
;  EARLIER APPLICATION NUMBER: PCT/US98/04493
;  EARLIER FILING DATE: 1998-03-06
;  EARLIER APPLICATION NUMBER: 60/040,162
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,333
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/038,621
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,626
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,334
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,336
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,163
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/047,600
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,615
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,597
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,502
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,633
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,583
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,617
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,618
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,503
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,592
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,581
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,584
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,500
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,587
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,492
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,598
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,613
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,582
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,596
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,612
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,632
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,601
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/043,580
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,568
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,314
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,569
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,311
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,671
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,674
;  EARLIER FILING DATE: 1997-04-11
```

; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match 9.6%; Score 30.8; DB 4; Length 1695;
Best Local Similarity 58.9%; Pred. No. 6.8;
Matches 53; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```
Qy    111 AGAGGTGGCTGTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGAT 170
          ||||||||| ||||    || || |  || || || |  | ||||  || || |  |   |
Db   1382 AGAGGTGGATGTTTGGGATTGTATTATTATTTTTACTAATATCTGTAGCTATTTTGTTTTT 1441

Qy    171 TCCTGTATTTGTAGTTAAAATCCTTTCCCT 200
          | || |  || | |||    ||| ||  ||
Db   1442 TGCTTTGGTTATTGTTTTTTTCCCTTTTCT 1471
```

RESULT 8
US-08-947-823-1/c
; Sequence 1, Application US/08947823
; Patent No. 6114605
; GENERAL INFORMATION:
; APPLICANT: Williamson, Valerie M.
; APPLICANT: Kaloshian, Isgouhi
; APPLICANT: Yaghoobi, Jafar
; APPLICANT: Bodeau, John
; APPLICANT: Milligan, Stephen
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Pest Resistance in Plants
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:

```
;       ADDRESSEE:  Townsend and Townsend and Crew LLP
;       STREET:  Two Embarcadero Center, Eighth Floor
;       CITY:  San Francisco
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  94111-3834
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/947,823
;       FILING DATE:  09-OCT-1997
;       CLASSIFICATION:  800
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US97/18802
;       FILING DATE:  09-OCT-1997
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 60/028,191
;       FILING DATE:  10-OCT-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Bastian, Kevin L.
;       REGISTRATION NUMBER:  34,774
;       REFERENCE/DOCKET NUMBER:  02307O-070210US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 576-0200
;       TELEFAX:  (415) 576-0300
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  51952 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-08-947-823-1

  Query Match             9.5%;  Score 30.6;  DB 3;  Length 51952;
  Best Local Similarity  56.4%;  Pred. No. 22;
  Matches   57;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy      91 GTAAGGAATTGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTTATAGTGTTC 150
           |||   ||||  |||        ||  | | ||||    |||| |   ||||||||
Db  23249 GTAGAAAATTTCTTTTTTGTGTAGTCGCCAGTTGATTTTCGTATGAGTTTTATAGGCAGA 23190

Qy     151 TCTGATGCCATCTGCTGGATTCCTGTATTTGTAGTTAAAAT 191
           | ||  || | |||||  |    ||| | |. ||| ||||||
Db  23189 TGTGCGGCTAGCTGCTTAAACTTTGTTTGTTTAGGTAAAAT 23149


RESULT 9
US-08-475-886-5
; Sequence 5, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
;  APPLICANT: FUNKHOUSER, ANN W
;  APPLICANT: EMERSON, SUZANNE U
;  APPLICANT: PURCELL, ROBERT H
;  APPLICANT: D'HONDT, ERIC
;  TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
;  FILE REFERENCE: 20264262US2
;  CURRENT APPLICATION NUMBER: US/08/475,886A
;  CURRENT FILING DATE: 1995-06-07
;  EARLIER APPLICATION NUMBER: 07/947,338
;  EARLIER FILING DATE: 1992-09-18
;  EARLIER APPLICATION NUMBER: 08/397,232
;  EARLIER FILING DATE: 1995-03-10
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 7486
;   TYPE: DNA
;   ORGANISM: Attenuated HAV (4380), strain HM-175
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (730)..(7410)
US-08-475-886-5

  Query Match             9.4%;  Score 30.2;  DB 3;  Length 7486;
  Best Local Similarity  47.6%;  Pred. No. 16;
  Matches   89;  Conservative    0;  Mismatches   98;  Indels    0;  Gaps    0;

Qy      40 TATATTACTATGTTCTGTTCCATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAAT 99
           | ||||| | ||||| || |  ||   | ||     |   |||  |||||  |||    |
Db   7298 TTTATTATTTTGTTCAGTCCTGTTTGGAGAAAGAGATGATAGAATACAGACTTAAATCTT 7357

Qy     100 TGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCC 159
             ||||  ||||    | | ||   |   || | |  ||| |    ||   ||||
Db   7358 ATGATTGGTGGAGAATGAGATTTTATGACCAGTGTTTCATTTGTGACCTTTCATGATTTG 7417

Qy     160 ATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCA 219
           |      | |||    ||    |  | |   |  |||   ||||      | |||||  |
Db   7418 TTTAAACGAATTTTCTTAAAATTTCTGAGGTTTGTTTATTTCTTTTATCAGTAAATAAAA 7477

Qy     220 GACACAA 226
           | | ||
Db   7478 AAAAAAA 7484


RESULT 10
US-08-397-232-3
; Sequence 3, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
;  APPLICANT: FUNKHOUSER, ANN W
;  APPLICANT: EMERSON, SUZANNE U
;  APPLICANT: PURCELL, ROBERT H
;  APPLICANT: D'HONDT, ERIC
;  TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
;  FILE REFERENCE: 20264262US1
;  CURRENT APPLICATION NUMBER: US/08/397,232A
;  CURRENT FILING DATE: 1995-04-17
;  EARLIER APPLICATION NUMBER: 07/947,338
;  EARLIER FILING DATE: 1992-09-18
;  EARLIER APPLICATION NUMBER: PCT/US93/08610
;  EARLIER FILING DATE: 1993-09-17
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 7486
;   TYPE: DNA
;   ORGANISM: Attenuated HAV (4380), strain HM-175
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (730)..(7410)
US-08-397-232-3

  Query Match             9.4%;  Score 30.2;  DB 4;  Length 7486;
  Best Local Similarity  47.6%;  Pred. No. 16;
  Matches   89;  Conservative    0;  Mismatches   98;  Indels    0;  Gaps    0;

Qy      40 TATATTACTATGTTCTGTTCCATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAAT 99
           | ||||| | ||||| || |  ||   | ||     |   |||  |||||  |||    |
Db   7298 TTTATTATTTTGTTCAGTCCTGTTTGGAGAAAGAGATGATAGAATACAGACTTAAATCTT 7357

Qy     100 TGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCC 159
             ||||  ||||    | | ||   |   || | |  ||| |    ||   ||||
Db   7358 ATGATTGGTGGAGAATGAGATTTTATGACCAGTGTTTCATTTGTGACCTTTCATGATTTG 7417

Qy     160 ATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCA 219
           |      | |||    ||    |  | |   |  |||   ||||      | |||||  |
Db   7418 TTTAAACGAATTTTCTTAAAATTTCTGAGGTTTGTTTATTTCTTTTATCAGTAAATAAAA 7477

Qy     220 GACACAA 226
           | | ||
Db   7478 AAAAAAA 7484
```

```
RESULT 11
US-09-653-499-5
; Sequence 5, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
;  APPLICANT: FUNKHOUSER, ANN W
;  APPLICANT:  EMERSON, SUZANNE U
;  APPLICANT:  PURCELL, ROBERT H
;  APPLICANT:  D'HONDT, ERIC
;  TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
;  FILE REFERENCE: 20264262US2
;  CURRENT APPLICATION NUMBER: US/09/653,499
;  CURRENT FILING DATE:  2000-08-31
;  PRIOR APPLICATION NUMBER: 08/475,886
;  PRIOR FILING DATE: 1999-08-09
;  PRIOR APPLICATION NUMBER: 07/947,338
;  PRIOR FILING DATE: 1992-09-18
;  PRIOR APPLICATION NUMBER: 08/397,232
;  PRIOR FILING DATE: 1995-03-10
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 7486
;   TYPE: DNA
;   ORGANISM: Attenuated HAV (4380), strain HM-175
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (730)..(7410)
US-09-653-499-5

  Query Match             9.4%;  Score 30.2;  DB 4;  Length 7486;
  Best Local Similarity   47.6%;  Pred. No. 16;
  Matches   89;  Conservative    0;  Mismatches   98;  Indels    0;  Gaps    0;

Qy      40 TATATTACTATGTTCTGTTCCATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAAT 99
           | ||||| | ||||| || |  ||   | ||     |   |||  |||||  |||    |
Db    7298 TTTATTATTTTGTTCAGTCCTGTTTGGAGAAAGAGATGATAGAATACAGACTTAAATCTT 7357

Qy     100 TGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCC 159
             ||||  ||||    | | ||   |   || | |  ||| |    ||   ||||
Db    7358 ATGATTGGTGGAGAATGAGATTTTATGACCAGTGTTTCATTTGTGACCTTTCATGATTTG 7417

Qy     160 ATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCA 219
           |     | |||    ||    |  | |   |  |||   ||||      | |||||  |
Db    7418 TTTAAACGAATTTTCTTAAAATTTCTGAGGTTTGTTTATTTCTTTTATCAGTAAATAAAA 7477

Qy     220 GACACAA 226
           | | ||
Db    7478 AAAAAAA 7484


RESULT 12
US-08-475-886-3
; Sequence 3, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
;  APPLICANT: FUNKHOUSER, ANN W
;  APPLICANT: EMERSON, SUZANNE U
;  APPLICANT: PURCELL, ROBERT H
;  APPLICANT: D'HONDT, ERIC
;  TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
;  FILE REFERENCE: 20264262US2
;  CURRENT APPLICATION NUMBER: US/08/475,886A
;  CURRENT FILING DATE: 1995-06-07
;  EARLIER APPLICATION NUMBER: 07/947,338
;  EARLIER FILING DATE: 1992-09-18
;  EARLIER APPLICATION NUMBER: 08/397,232
;  EARLIER FILING DATE: 1995-03-10
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 7488
;   TYPE: DNA
;   ORGANISM: Attenuated (Pass 35) HAV, strain HM-175
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (730)..(7410)
US-08-475-886-3

  Query Match             9.4%;  Score 30.2;  DB 3;  Length 7488;
  Best Local Similarity   47.6%;  Pred. No. 16;
  Matches   89;  Conservative    0;  Mismatches   98;  Indels    0;  Gaps    0;

Qy      40 TATATTACTATGTTCTGTTCCATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAAT 99
           | ||||| | ||||| || |  ||   | ||     |   |||  |||||  |||    |
Db    7298 TTTATTATTTTGTTCAGTCCTGTTTGGAGAAAGAGATGATAGAATACAGACTTAAATCTT 7357

Qy     100 TGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCC 159
             ||||  ||||    | | ||   |   || | |  ||| |    ||   ||||
Db    7358 ATGATTGGTGGAGAATGAGATTTTATGACCAGTGTTTCATTTGTGACCTTTCATGATTTG 7417

Qy     160 ATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCA 219
           |     | |||    ||    |  | |   |  |||   ||||      | |||||  |
Db    7418 TTTAAACGAATTTTCTTAAAATTTCTGAGGTTTGTTTATTTCTTTTATCAGTAAATAAAA 7477

Qy     220 GACACAA 226
           | | ||
Db    7478 AAAAAAA 7484


RESULT 13
US-09-653-499-3
; Sequence 3, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
;  APPLICANT: FUNKHOUSER, ANN W
;  APPLICANT:  EMERSON, SUZANNE U
;  APPLICANT:  PURCELL, ROBERT H
;  APPLICANT:  D'HONDT, ERIC
;  TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
;  FILE REFERENCE: 20264262US2
;  CURRENT APPLICATION NUMBER: US/09/653,499
;  CURRENT FILING DATE:  2000-08-31
;  PRIOR APPLICATION NUMBER: 08/475,886
;  PRIOR FILING DATE: 1999-08-09
;  PRIOR APPLICATION NUMBER: 07/947,338
;  PRIOR FILING DATE: 1992-09-18
;  PRIOR APPLICATION NUMBER: 08/397,232
;  PRIOR FILING DATE: 1995-03-10
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 7488
;   TYPE: DNA
;   ORGANISM: Attenuated (Pass 35) HAV, strain HM-175
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (730)..(7410)
US-09-653-499-3

  Query Match             9.4%;  Score 30.2;  DB 4;  Length 7488;
  Best Local Similarity   47.6%;  Pred. No. 16;
  Matches   89;  Conservative    0;  Mismatches   98;  Indels    0;  Gaps    0;

Qy      40 TATATTACTATGTTCTGTTCCATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAAT 99
           | ||||| | ||||| || |  ||   | ||     |   |||  |||||  |||    |
Db    7298 TTTATTATTTTGTTCAGTCCTGTTTGGAGAAAGAGATGATAGAATACAGACTTAAATCTT 7357

Qy     100 TGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCC 159
             ||||  ||||    | | ||   |   || | |  ||| |    ||   ||||
Db    7358 ATGATTGGTGGAGAATGAGATTTTATGACCAGTGTTTCATTTGTGACCTTTCATGATTTG 7417

Qy     160 ATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCA 219
           |     | |||    ||    |  | |   |  |||   ||||      | |||||  |
```

```
Db   7418 TTTAAACGAATTTTCTTAAAATTTCTGAGGTTTGTTTATTTCTTTTATCAGTAAATAAAA 7477

Qy    220 GACACAA 226
          | | ||
Db   7478 AAAAAAA 7484


RESULT 14
US-08-945-296-2
; Sequence 2, Application US/08945296
; Patent No. 5972609
;  GENERAL INFORMATION:
;    APPLICANT:  Tinsley, Jonathon M
;    APPLICANT:  Davies, Kay E
;    APPLICANT:  Dennis, Carina
;    TITLE OF INVENTION:  Utrophin gene promoter
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nixon & Vanderhye P.C.
;      STREET:  1100 No. 5972609th Glebe Road, 8th Floor
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  United States of America
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/945,296
;      FILING DATE:  22-OCT-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB96/00979
;      FILING DATE:  24-APR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9508236.8
;      FILING DATE:  24-APR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9526028.7
;      FILING DATE:  20-DEC-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ms Mary J Wilson
;      REGISTRATION NUMBER:  32,955
;      REFERENCE/DOCKET NUMBER:  620-26
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 816-4000
;      TELEFAX:  (703) 816-4100
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1655 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-945-296-2

  Query Match             9.3%;  Score 30;  DB 2;  Length 1655;
  Best Local Similarity   49.1%;  Pred. No. 11;
  Matches  108;  Conservative    0;  Mismatches  110;  Indels    2;  Gaps    1;

Qy      7 AACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCCATTCAA 66
          || ||| || ||   | |  | |   | ||| |  | | ||| | || | ||      ||
Db   1077 AAGTTGGTGCCTGCGCGCCCCTTCCAGGTTTGCGCTTTGACTGTTTTGTTTTTGGCGGAA 1136

Qy     67 AAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT--GTTG 124
            | | | |  | |||   || ||||||||     |||  |  | | |     |   |
Db   1137 CTACCAGGCAGGAAGATTGCACAAGTAAGGGGCGTTTTCAGTCGGGTGTCAATTTCTCTT 1196

Qy    125 CAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTTGTAG 184
            || ||  |||||| | | |    |  || |  || ||  |  | |   ||| |
Db   1197 TCTTTCTTTCTTTTTTTAAAATTTCGGTTCGTGTCTGCTTCTCCAAGCTTTATTTTTTTT 1256

Qy    185 TTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACAC 224
          |||||||  | |  |  | |    |   | |  || ||||
Db   1257 TTAAAATACATCGCACCACCAAACTAACACTCGCACACAC 1296


RESULT 15
US-09-405-112-2
; Sequence 2, Application US/09405112
; Patent No. 6087111
;  GENERAL INFORMATION:
;    APPLICANT:  Tinsley, Jonathon M
;    APPLICANT:  Davies, Kay E
;    APPLICANT:  Dennis, Carina
;    TITLE OF INVENTION:  Utrophin gene promoter
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nixon & Vanderhye P.C.
;      STREET:  1100 No. 6087111th Glebe Road, 8th Floor
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  United States of America
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/405,112
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/945,296
;      FILING DATE:  22-OCT-1997
;      APPLICATION NUMBER:  PCT/GB96/00979
;      FILING DATE:  24-APR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9508236.8
;      FILING DATE:  24-APR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9526028.7
;      FILING DATE:  20-DEC-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ms Mary J Wilson
;      REGISTRATION NUMBER:  32,955
;      REFERENCE/DOCKET NUMBER:  620-26
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 816-4000
;      TELEFAX:  (703) 816-4100
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1655 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-09-405-112-2

  Query Match             9.3%;  Score 30;  DB 3;  Length 1655;
  Best Local Similarity   49.1%;  Pred. No. 11;
  Matches  108;  Conservative    0;  Mismatches  110;  Indels    2;  Gaps    1;

Qy      7 AACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCCATTCAA 66
          || ||| || ||   | |  | |   | ||| |  | | ||| | || | ||      ||
Db   1077 AAGTTGGTGCCTGCGCGCCCCTTCCAGGTTTGCGCTTTGACTGTTTTGTTTTTGGCGGAA 1136

Qy     67 AAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT--GTTG 124
            | | | |  | |||   || ||||||||     |||  |  | | |     |   |
Db   1137 CTACCAGGCAGGAAGATTGCACAAGTAAGGGGCGTTTTCAGTCGGGTGTCAATTTCTCTT 1196

Qy    125 CAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTTGTAG 184
            || ||  |||||| | | |    |  || |  || ||  |  | |   ||| |
Db   1197 TCTTTCTTTCTTTTTTTAAAATTTCGGTTCGTGTCTGCTTCTCCAAGCTTTATTTTTTTT 1256

Qy    185 TTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACAC 224
          |||||||  | |  |  | |    |   | |  || ||||
```

Db 1257 TTAAAATACATCGCACCACCAAACTAACACTCGCACACAC 1296

Search completed: May 26, 2003, 20:05:41
Job time : 110 secs

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 18:59:22 ; Search time 133 Seconds
(without alignments)
3186.980 Million cell updates/sec

Title: US-09-930-312-1
Perfect score: 321
Sequence: 1 ggtgtgaacttgctggcttt.....acttttttaaggacaagttg 321

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

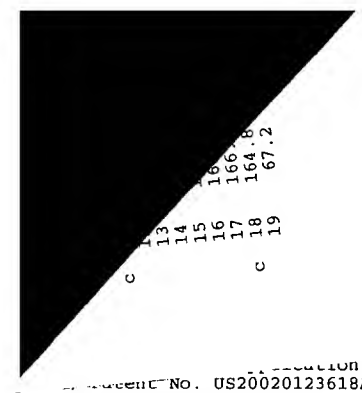
Result No.	Score	% Query Match	Length	DB	ID	Description
1	321	100.0	321	9	US-09-930-312-1	Sequence 1, Appli
2	321	100.0	2049	10	US-09-928-175-11	Sequence 11, Appl
3	321	100.0	2142	9	US-09-965-536A-5	Sequence 5, Appli
4	321	100.0	2193	10	US-09-928-175-6	Sequence 6, Appli
5	321	100.0	2214	9	US-09-965-536A-1	Sequence 1, Appli
6	321	100.0	2265	10	US-09-928-175-1	Sequence 1, Appli
7	321	100.0	2436	9	US-10-229-735-2	Sequence 2, Appli
8	321	100.0	2838	9	US-10-222-668-1	Sequence 1, Appli
9	233	72.6	2214	10	US-09-928-175-19	Sequence 19, Appl
10	233	72.6	2539	9	US-10-229-735-4	Sequence 4, Appli
11	220	68.5	1018	9	US-09-782-974C-65	Sequence 65, Appl
12	166.4	51.8	530	9	US-09-989-442-33	Sequence 33, Appl
13	166.4	51.8	530	9	US-10-073-865-25	Sequence 25, Appl
14	166.4	51.8	530	9	US-10-103-313-89	Sequence 89, Appl
15	166.4	51.8	530	10	US-09-764-853-137	Sequence 137, App
16	166.4	51.8	530	10	US-09-764-877-986	Sequence 986, App
17	166.4	51.8	1804	10	US-09-895-686-10	Sequence 10, Appl
18	166.4	51.3	612	10	US-09-895-686-37	Sequence 37, Appl
19	166.4	20.9	1015	9	US-10-270-333-20	Sequence 20, Appl
c 20	58	18.1	432	10	US-09-895-686-62	Sequence 62, Appl
21	53.2	16.6	1080	9	US-10-270-333-26	Sequence 26, Appl
22	53.2	16.6	3187	9	US-10-270-333-25	Sequence 25, Appl
23	39.2	12.2	3574	9	US-10-270-333-19	Sequence 19, Appl
24	34.4	10.7	2586	10	US-09-804-551B-19	Sequence 19, Appl
c 25	33	10.3	746	9	US-09-764-891-1631	Sequence 1631, Ap
26	32.8	10.2	3007	9	US-10-239-676-219	Sequence 219, App
c 27	32.8	10.2	3390	9	US-10-140-164-1	Sequence 1, Appli
c 28	32.8	10.2	3556	9	US-10-140-164-3	Sequence 3, Appli
29	32.4	10.1	1739	9	US-09-800-274-3	Sequence 3, Appli
c 30	32.4	10.1	32177	9	US-09-764-891-6967	Sequence 6967, Ap
c 31	32.4	10.1	32177	10	US-09-764-877-3251	Sequence 3251, Ap
32	32.2	10.0	7544	9	US-10-239-676-5	Sequence 5, Appli
c 33	31.6	9.8	478	9	US-09-918-995-12292	Sequence 12292, A
c 34	31.6	9.8	11036	9	US-10-239-676-117	Sequence 117, App
c 35	31.2	9.7	554	10	US-09-864-761-12015	Sequence 12015, A
c 36	31	9.7	15734	9	US-10-091-504-1922	Sequence 1922, Ap
c 37	31	9.7	15734	10	US-09-764-869-1922	Sequence 1922, Ap
38	30.8	9.6	1472	9	US-09-746-783-185	Sequence 185, App
39	30.8	9.6	1695	9	US-09-809-391-299	Sequence 299, App
c 40	30.8	9.6	2000	9	US-09-938-842A-3915	Sequence 3915, Ap
41	30.8	9.6	172637	10	US-09-805-458A-3	Sequence 3, Appli
c 42	30.6	9.5	714	9	US-09-935-338-232	Sequence 232, App
43	30.6	9.5	918	9	US-09-991-936-1572	Sequence 1572, Ap
44	30.6	9.5	2965	10	US-09-764-864-278	Sequence 278, App
45	30.6	9.5	2970	10	US-09-764-864-208	Sequence 208, App

ALIGNMENTS

RESULT 1
US-09-930-312-1
; Sequence 1, Application US/09930312
; Publication No. US20030032019A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn Company
; APPLICANT: Lind, Peter
; TITLE OF INVENTION: No. US20030032019Alel G Protein-Coupled Receptors
; FILE REFERENCE: 00329.PCT1
; CURRENT APPLICATION NUMBER: US/09/930,312
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: USSN 60/225,262
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 321
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-312-1

Query Match 100.0%; Score 321; DB 9; Length 321;
Best Local Similarity 100.0%; Pred. No. 2.5e-85;
Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
```



```
                                                ||||||||||||||||||||||||||||||||||||||||
                                                GTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300

                                                TTG 321
                                                |||
                                                TTG 321
```

...plication US/09928175
...ent No. US20020123618A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher J.
; APPLICANT: Gong, Jianhua
; APPLICANT: Daugherty, Betsy
; APPLICANT: Rogers, No. US20020123618A1ma
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 00-1229
; CURRENT APPLICATION NUMBER: US/09/928,175
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/224,455
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2049
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2046)
; NAME/KEY: sig_peptide
; LOCATION: (1)..(108)
US-09-928-175-11

Query Match 100.0%; Score 321; DB 10; Length 2049;
Best Local Similarity 100.0%; Pred. No. 5.9e-85;
Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1570 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1629

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1630 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1689

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1690 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1749

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1750 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1809

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1810 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 1869

Qy     301 AACTTTTTTAAGGACAAGTTG 321
          |||||||||||||||||||||
Db    1870 AACTTTTTTAAGGACAAGTTG 1890
```

RESULT 3
US-09-965-536A-5
; Sequence 5, Application US/09965536A
; Publication No. US20030027323A1
; GENERAL INFORMATION:
; APPLICANT: FEDER, J. N.
; APPLICANT: MINTIER, G.
; APPLICANT: RAMANATHAN, C. S.
; APPLICANT: HAWKEN, D. R.
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMY5,
; TITLE OF INVENTION: EXPRESSED HIGHLY IN BRAIN AND OVARIAN TISSUES
; FILE REFERENCE: DO041NP
; CURRENT APPLICATION NUMBER: US/09/965,536A
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/235,713
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/261,781
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/306,605
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/310,436
; PRIOR FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 2142
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-965-536A-5

Query Match 100.0%; Score 321; DB 9; Length 2142;
Best Local Similarity 100.0%; Pred. No. 6e-85;
Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1663 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1722

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1723 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1782

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1783 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1842

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1843 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1902

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1903 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 1962

Qy     301 AACTTTTTTAAGGACAAGTTG 321
          |||||||||||||||||||||
Db    1963 AACTTTTTTAAGGACAAGTTG 1983
```

RESULT 4
US-09-928-175-6
; Sequence 6, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher J.
; APPLICANT: Gong, Jianhua
; APPLICANT: Daugherty, Betsy
; APPLICANT: Rogers, No. US20020123618A1ma
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 00-1229
; CURRENT APPLICATION NUMBER: US/09/928,175
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/224,455
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 2193
; TYPE: DNA

```
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(2190)
;   NAME/KEY: sig_peptide
;   LOCATION: (1)..(108)
US-09-928-175-6

  Query Match            100.0%;  Score 321;  DB 10;  Length 2193;
  Best Local Similarity  100.0%;  Pred. No. 6.1e-85;
  Matches  321;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1714 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1773

Qy     61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1774 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1833

Qy    121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1834 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1893

Qy    181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1894 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1953

Qy    241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1954 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2013

Qy    301 AACTTTTTTAAGGACAAGTTG 321
          |||||||||||||||||||||
Db   2014 AACTTTTTTAAGGACAAGTTG 2034


RESULT 5
US-09-965-536A-1
; Sequence 1, Application US/09965536A
; Publication No. US20030027323A1
; GENERAL INFORMATION:
;  APPLICANT: FEDER, J. N.
;  APPLICANT:  MINTIER, G.
;  APPLICANT:  RAMANATHAN, C. S.
;  APPLICANT:  HAWKEN, D. R.
;  TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMY5,
;  TITLE OF INVENTION:  EXPRESSED HIGHLY IN BRAIN AND OVARIAN TISSUES
;  FILE REFERENCE: DO041NP
;  CURRENT APPLICATION NUMBER: US/09/965,536A
;  CURRENT FILING DATE:  2001-09-26
;  PRIOR APPLICATION NUMBER: 60/235,713
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: 60/261,781
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: 60/306,605
;  PRIOR FILING DATE: 2001-07-19
;  PRIOR APPLICATION NUMBER: 60/310,436
;  PRIOR FILING DATE: 2001-08-03
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 2214
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-965-536A-1

  Query Match            100.0%;  Score 321;  DB 9;  Length 2214;
  Best Local Similarity  100.0%;  Pred. No. 6.1e-85;
  Matches  321;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1735 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1794

Qy     61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1795 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1854

Qy    121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1855 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1914

Qy    181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1915 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1974

Qy    241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1975 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2034

Qy    301 AACTTTTTTAAGGACAAGTTG 321
          |||||||||||||||||||||
Db   2035 AACTTTTTTAAGGACAAGTTG 2055


RESULT 6
US-09-928-175-1
; Sequence 1, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
;  APPLICANT: Paszty, Christopher J.
;  APPLICANT:  Gong, Jianhua
;  APPLICANT:  Daugherty, Betsy
;  APPLICANT:  Rogers, No. US20020123618A1ma
;  TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
;  TITLE OF INVENTION:  Uses Thereof
;  FILE REFERENCE: 00-1229
;  CURRENT APPLICATION NUMBER: US/09/928,175
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: 60/224,455
;  PRIOR FILING DATE: 2000-08-10
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 2265
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(2262)
;   NAME/KEY: sig_peptide
;   LOCATION: (1)..(108)
US-09-928-175-1

  Query Match            100.0%;  Score 321;  DB 10;  Length 2265;
  Best Local Similarity  100.0%;  Pred. No. 6.2e-85;
  Matches  321;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1786 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1845

Qy     61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1846 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1905

Qy    121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1906 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1965

Qy    181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1966 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 2025

Qy    241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2026 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2085

Qy     301 AACTTTTTTAAGGACAAGTTG 321
           |||||||||||||||||||||
Db    2086 AACTTTTTTAAGGACAAGTTG 2106


RESULT 7
US-10-229-735-2
; Sequence 2, Application US/10229735
; Publication No. US20030082650A1
; GENERAL INFORMATION:
;  APPLICANT: Baylor College of Medicine
;  APPLICANT:  Agoulnik, Alexander I.
;  TITLE OF INVENTION: The GREAT Gene and Protein
;  FILE REFERENCE: 7572/73263
;  CURRENT APPLICATION NUMBER: US/10/229,735
;  CURRENT FILING DATE:  2002-08-29
;  PRIOR APPLICATION NUMBER: 60/315,696
;  PRIOR FILING DATE: 2001-08-30
;  PRIOR APPLICATION NUMBER: 60/351,432
;  PRIOR FILING DATE: 2002-01-28
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 2436
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-229-735-2

  Query Match             100.0%;  Score 321;  DB 9;  Length 2436;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-85;
  Matches  321;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1856 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1915

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1916 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1975

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1976 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 2035

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2036 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 2095

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2096 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2155

Qy     301 AACTTTTTTAAGGACAAGTTG 321
           |||||||||||||||||||||
Db    2156 AACTTTTTTAAGGACAAGTTG 2176


RESULT 8
US-10-222-668-1
; Sequence 1, Application US/10222668
; Publication No. US20030088884A1
; GENERAL INFORMATION:
;  APPLICANT: Hsu, Sheau Yu
;  APPLICANT:  Hsueh, Aaron
;  TITLE OF INVENTION: Mammalian Relaxin Receptor
;  FILE REFERENCE: STAN-239 WO
;  CURRENT APPLICATION NUMBER: US/10/222,668
;  CURRENT FILING DATE:  2002-08-15
;  PRIOR APPLICATION NUMBER: 60/313,259
;  PRIOR FILING DATE: 2002-08-17
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 2838
;   TYPE: DNA
;   ORGANISM: H. sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (107)...(2369)
US-10-222-668-1

  Query Match             100.0%;  Score 321;  DB 9;  Length 2838;
  Best Local Similarity   100.0%;  Pred. No. 6.9e-85;
  Matches  321;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1892 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1951

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1952 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 2011

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2012 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 2071

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2072 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 2131

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2132 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2191

Qy     301 AACTTTTTTAAGGACAAGTTG 321
           |||||||||||||||||||||
Db    2192 AACTTTTTTAAGGACAAGTTG 2212


RESULT 9
US-09-928-175-19
; Sequence 19, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
;  APPLICANT: Paszty, Christopher J.
;  APPLICANT:  Gong, Jianhua
;  APPLICANT:  Daugherty, Betsy
;  APPLICANT:  Rogers, No. US20020123618A1ma
;  TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
;  TITLE OF INVENTION:  Uses Thereof
;  FILE REFERENCE: 00-1229
;  CURRENT APPLICATION NUMBER: US/09/928,175
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: 60/224,455
;  PRIOR FILING DATE: 2000-08-10
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
;   LENGTH: 2214
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(2211)
;   NAME/KEY: sig_peptide
;   LOCATION: (1)..(57)
US-09-928-175-19

  Query Match             72.6%;  Score 233;  DB 10;  Length 2214;
  Best Local Similarity   82.9%;  Pred. No. 6.9e-59;
  Matches  266;  Conservative    0;  Mismatches   55;  Indels    0;  Gaps    0;
```

```
Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
           ||||||||||||||||||| ||| |||| ||| |||||||| | || |||||||| |||
Db    1735 GGTGTGAACTTGCTGGCTTTCCTCGTCATCGTGATTTCCTATGTCACCATGTTCTGCTCC 1794

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
           ||||| ||||| ||| | ||||| |||||||| |||| | || | |||||||||
Db    1795 ATTCATAAAACAGCCCTTCAGACTGCAGAAGTGAGGAGCCACATCGGGAAGGAGGTGGCT 1854

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
           |||||||| || |||||||||| ||||||||||||||||||||||||| ||||| |||
Db    1855 GTTGCAAACCGGTTCTTTTTTATCGTGTTCTCTGATGCCATCTGCTGGATCCCTGTGTTT 1914

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
           || ||||| ||||| || ||| | |  ||||| ||||||| |||||| ||||||||||
Db    1915 GTCGTTAAGATCCTGTCTCTCCTTCAAGTGGAGATACCAGGCACAATCACTTCCTGGATC 1974

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
           ||| |||||||||||| || ||||| || || || || |||||||| ||||| || |||
Db    1975 GTGGTTTTTTTCCTTCCGGTGAACAGCGCCTTAAACCCCATCCTCTACACTCTGACGACC 2034

Qy     301 AACTTTTTTAAGGACAAGTTG 321
           |||||||||||||||||||
Db    2035 TCCTTTTTTAAGGACAAGTTG 2055


RESULT 10
US-10-229-735-4
; Sequence 4, Application US/10229735
; Publication No. US20030082650A1
; GENERAL INFORMATION:
;  APPLICANT: Baylor College of Medicine
;  APPLICANT:  Agoulnik, Alexander I.
;  TITLE OF INVENTION: The GREAT Gene and Protein
;  FILE REFERENCE: 7572/73263
;  CURRENT APPLICATION NUMBER: US/10/229,735
;  CURRENT FILING DATE:  2002-08-29
;  PRIOR APPLICATION NUMBER: 60/315,696
;  PRIOR FILING DATE: 2001-08-30
;  PRIOR APPLICATION NUMBER: 60/351,432
;  PRIOR FILING DATE: 2002-01-28
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 2539
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-229-735-4

  Query Match            72.6%;  Score 233;  DB 9;  Length 2539;
  Best Local Similarity  82.9%;  Pred. No. 7.3e-59;
  Matches  266;  Conservative    0;  Mismatches   55;  Indels    0;  Gaps    0;

Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
           ||||||||||||||||||| ||| |||| ||| |||||||| | || |||||||| |||
Db    1928 GGTGTGAACTTGCTGGCTTTCCTCGTCATCGTGATTTCCTATGTCACCATGTTCTGCTCC 1987

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
           ||||| ||||| ||| | ||||| |||||||| |||| | || | |||||||||
Db    1988 ATTCATAAAACAGCCCTTCAGACTGCAGAAGTGAGGAGCCACATCGGGAAGGAGGTGGCT 2047

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
           |||||||| || |||||||||| ||||||||||||||||||||||||| ||||| |||
Db    2048 GTTGCAAACCGGTTCTTTTTTATCGTGTTCTCTGATGCCATCTGCTGGATCCCTGTGTTT 2107

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
           || ||||| ||||| || ||| | |  ||||| ||||||| |||||| ||||||||||
Db    2108 GTCGTTAAGATCCTGTCTCTCCTTCAAGTGGAGATACCAGGCACAATCACTTCCTGGATC 2167

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
           ||| |||||||||||| || ||||| || || || || |||||||| ||||| || |||
Db    2168 GTGGTTTTTTTCCTTCCGGTGAACAGCGCCTTAAACCCCATCCTCTACACTCTGACGACC 2227

Qy     301 AACTTTTTTAAGGACAAGTTG 321
           |||||||||||||||||||
Db    2228 TCCTTTTTTAAGGACAAGTTG 2248


RESULT 11
US-09-782-974C-65/c
; Sequence 65, Application US/09782974C
; Publication No. US20030082534A1
; GENERAL INFORMATION:
;  APPLICANT: Vogeli, Gabriel
;  APPLICANT:  Lind, Peter
;  APPLICANT:  Wood, Linda S.
;  APPLICANT:  Parodi, Luis A.
;  TITLE OF INVENTION: No. US20030082534A1el G Protein Coupled Receptor
;  FILE REFERENCE: 411USPHRM311
;  CURRENT APPLICATION NUMBER: US/09/782,974C
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: 60/165,838
;  PRIOR FILING DATE: 1999-11-16
;  PRIOR APPLICATION NUMBER: 09/714,449
;  PRIOR FILING DATE: 2000-11-16
;  PRIOR APPLICATION NUMBER: 60/198,568
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: 60/166,071
;  PRIOR FILING DATE: 1999-11-17
;  PRIOR APPLICATION NUMBER: 60/166,678
;  PRIOR FILING DATE: 1999-11-19
;  PRIOR APPLICATION NUMBER: 60/173,396
;  PRIOR FILING DATE: 1999-12-28
;  PRIOR APPLICATION NUMBER: 60/184,129
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/185,421
;  PRIOR FILING DATE: 2000-02-28
;  PRIOR APPLICATION NUMBER: 60/185,554
;  PRIOR FILING DATE: 2000-02-28
;  PRIOR APPLICATION NUMBER: 60/186,530
;  PRIOR FILING DATE: 2000-03-02
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 192
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 65
;   LENGTH: 1018
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-782-974C-65

  Query Match            68.5%;  Score 220;  DB 9;  Length 1018;
  Best Local Similarity  100.0%;  Pred. No. 3.4e-55;
  Matches  220;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     539 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 480

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     479 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 420

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     419 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 360

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAG 220
           |||||||||||||||||||||||||||||||||||||||
Db     359 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAG 320


RESULT 12
US-09-989-442-33
; Sequence 33, Application US/09989442
; Publication No. US20030013649A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
```

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJZ08
; CURRENT APPLICATION NUMBER: US/09/989,442
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17

```
;  PRIOR APPLICATION NUMBER: 60/249,244
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,217
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,211
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,215
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,264
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,214
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,297
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/232,400
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/231,242
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/232,081
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/232,080
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/231,414
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/231,244
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/233,064
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/233,063
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/232,397
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/232,399
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/232,401
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/241,808
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/241,826
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/241,786
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/241,221
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/246,475
;  PRIOR FILING DATE: 2000-11-08
;  PRIOR APPLICATION NUMBER: 60/231,243
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/233,065
;  PRIOR FILING DATE: 2000-09-14

  Query Match           51.8%;  Score 166.4;  DB 9;  Length 530;
  Best Local Similarity  70.0%;  Pred. No. 1.8e-39;
  Matches  224;  Conservative    0;  Mismatches   96;  Indels    0;  Gaps    0;

Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
           ||| | || |||   || ||| |||||| || |||||||||   |  ||||| | |
Db      32 GGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGT 91

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
           ||||  |||  ||| |     | || ||| || |||||    ||  || |||| ||
Db      92 GTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATC 151

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
           |||| || ||||| || |||||||| ||  |||||||  | |||||||| ||  | |||
Db     152 CTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTT 211

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
           ||||| ||| | ||||| ||  | | ||| ||||||||||  || || || || ||| ||
Db     212 GTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTA 271

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
           |||||||||| | || ||  |||||||||||||||| ||||| |||||||||| || ||
Db     272 GTGATTTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACA 331

Qy     301 AACTTTTTTAAGGACAAGTT 320
           |     ||||| || | | |
Db     332 AGACCATTTAAAGAAATGAT 351


RESULT 13
US-10-073-865-25
; Sequence 25, Application US/10073865
; Publication No. US20030044904A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PJZ09C1
;  CURRENT APPLICATION NUMBER: US/10/073,865
;  CURRENT FILING DATE:  2002-02-14
;  Prior Application removed - See file Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 154
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
;   LENGTH: 530
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-073-865-25

  Query Match           51.8%;  Score 166.4;  DB 9;  Length 530;
  Best Local Similarity  70.0%;  Pred. No. 1.8e-39;
  Matches  224;  Conservative    0;  Mismatches   96;  Indels    0;  Gaps    0;

Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
           ||| | || |||   || ||| |||||| || |||||||||   |  ||||| | |
Db      32 GGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGT 91

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
           ||||  |||  ||| |     | || ||| || |||||    ||  || |||| ||
Db      92 GTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATC 151

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
           |||| || ||||| || |||||||| ||  |||||||  | |||||||| ||  | |||
Db     152 CTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTT 211

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
           ||||| ||| | ||||| ||  | | ||| ||||||||||  || || || || ||| ||
Db     212 GTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTA 271

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
           |||||||||| | || ||  |||||||||||||||| ||||| |||||||||| || ||
Db     272 GTGATTTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACA 331

Qy     301 AACTTTTTTAAGGACAAGTT 320
           |     ||||| || | | |
Db     332 AGACCATTTAAAGAAATGAT 351


RESULT 14
US-10-103-313-89
; Sequence 89, Application US/10103313
; Publication No. US20030082758A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PJZ07C1
;  CURRENT APPLICATION NUMBER: US/10/103,313
;  CURRENT FILING DATE:  2002-03-12
;  NUMBER OF SEQ ID NOS: 653
;       Prior Application removed - See File Wrapper or Palm
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 89
;   LENGTH: 530
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-103-313-89
```

```
  Query Match             51.8%;  Score 166.4;  DB 9;  Length 530;
  Best Local Similarity   70.0%;  Pred. No. 1.8e-39;
  Matches  224;  Conservative    0;  Mismatches   96;  Indels    0;  Gaps    0;

Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
           ||| | || |||   || ||| ||||||| || |||||||||   |  ||||| | |
Db      32 GGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGT 91

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
             ||||  |||  ||| |     | || ||| || |||||    ||  || |||| ||
Db      92 GTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATC 151

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
             |||| || ||||| || |||||||| ||  |||||||  | |||||||| ||  | |||
Db     152 CTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTT 211

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
           ||||| ||| | ||||| || | | ||| ||||||||||  || || || || ||| ||
Db     212 GTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTA 271

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
           |||||||| | || ||  |||||||||||||||| ||||| |||||||||| || ||
Db     272 GTGATTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACA 331

Qy     301 AACTTTTTTAAGGACAAGTT 320
           |     ||||| || | | |
Db     332 AGACCATTTAAAGAAATGAT 351


RESULT 15
US-09-764-853-137
; Sequence 137, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PJZ06
;  CURRENT APPLICATION NUMBER: US/09/764,853
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 939
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 137
;   LENGTH: 530
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-853-137

  Query Match             51.8%;  Score 166.4;  DB 10;  Length 530;
  Best Local Similarity   70.0%;  Pred. No. 1.8e-39;
  Matches  224;  Conservative    0;  Mismatches   96;  Indels    0;  Gaps    0;

Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
           ||| | || |||   || ||| ||||||| || |||||||||   |  ||||| | |
Db      32 GGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGT 91

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
             ||||  |||  ||| |     | || ||| || |||||    ||  || |||| ||
Db      92 GTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATC 151

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
             |||| || ||||| || |||||||| ||  |||||||  | |||||||| ||  | |||
Db     152 CTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTT 211

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
           ||||| ||| | ||||| || | | ||| ||||||||||  || || || || ||| ||
Db     212 GTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTA 271

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
           |||||||| | || ||  |||||||||||||||| ||||| |||||||||| || ||
Db     272 GTGATTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACA 331

Qy     301 AACTTTTTTAAGGACAAGTT 320
           |     ||||| || | | |
Db     332 AGACCATTTAAAGAAATGAT 351

Search completed: May 26, 2003, 20:07:59
Job time : 181 secs
```

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 18:40:32 ; Search time 1741 Seconds
(without alignments)
2986.072 Million cell updates/sec

Title: US-09-930-312-1
Perfect score: 321
Sequence: 1 ggtgtgaacttgctggcttt.....acttttttaaggacaagttg 321

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	186.2	58.0	197	12	BF942735	BF942735 EST-CD15N
	2	159	49.5	841	14	BQ228832	BQ228832 AGENCOURT
c	3	154.6	48.2	744	12	BG304121	BG304121 fl30f05.x
	4	152	47.4	235	10	AW436170	AW436170 75600 MAR
	5	77.4	24.1	456	12	BF021857	BF021857 uy58h08.y
	6	77.4	24.1	574	10	BB632910	BB632910 BB632910
c	7	43.2	13.5	1100	17	CNS016W7	AL107281 Drosophil
	8	40.6	12.6	992	17	CNS04L3O	AL295629 Tetraodon
	9	40	12.5	660	17	BH446416	BH446416 BOHFU39TR
c	10	39.2	12.2	1101	17	CNS00BEG	AL056872 Drosophil
	11	39	12.1	1024	17	CNS015ZE	AL106100 Drosophil
	12	38.2	11.9	681	17	AG148351	AG148351 Pan trogl
c	13	38	11.8	1073	9	AL551397	AL551397 AL551397
c	14	37.8	11.8	598	17	FR0020632	AL013515 F.rubripe
	15	37.8	11.8	1201	17	CNS016CF	AL106569 Drosophil
	16	37.6	11.7	1101	17	CNS0039G	AL063921 Drosophil
c	17	37.4	11.7	312	12	BF368833	BF368833 RC1-GN007
c	18	37.4	11.7	480	9	AJ460181	AJ460181 AJ460181
	19	37.4	11.7	624	10	AV929568	AV929568 AV929568
c	20	37.4	11.7	629	14	BQ465627	BQ465627 HU04D09r
c	21	37.2	11.6	441	9	AL369230	AL369230 MtBA29F09
	22	37.2	11.6	537	14	BM780117	BM780117 EST590693
c	23	37.2	11.6	654	10	BE239472	BE239472 EST403521
	24	37.2	11.6	887	14	BM779689	BM779689 EST590265
c	25	37	11.5	1092	17	CNS020K7	AL175696 Tetraodon
c	26	37	11.5	1433	12	BG247718	BG247718 602359413
c	27	36.8	11.5	619	17	FR0020635	AL013518 F.rubripe
c	28	36.8	11.5	678	10	AW984751	AW984751 RC1-HN001
	29	36.6	11.4	1201	17	CNS0107Y	AL098632 Drosophil
c	30	36.4	11.3	595	13	BM370473	BM370473 EBro08_SQ
	31	36.4	11.3	877	12	BF208790	BF208790 601872268
	32	36.4	11.3	949	17	CNS073SC	AL427906 clone BA0
c	33	36.2	11.3	267	9	AA993976	AA993976 ou34d12.s
	34	36.2	11.3	586	10	AV944500	AV944500 AV944500
	35	36.2	11.3	621	13	BJ484370	BJ484370 BJ484370
c	36	36	11.2	538	13	BJ062617	BJ062617 BJ062617
c	37	36	11.2	624	13	BJ060185	BJ060185 BJ060185
	38	35.8	11.2	507	14	BQ313904	BQ313904 QV3-BN014
c	39	35.8	11.2	515	9	AL789330	AL789330 AL789330
c	40	35.8	11.2	797	17	BH122046	BH122046 RPCI-24-3
	41	35.8	11.2	902	17	CNS00CVW	AL060007 Drosophil
c	42	35.6	11.1	360	9	AJ460183	AJ460183 AJ460183
	43	35.6	11.1	519	14	BM960895	BM960895 ciha9E13S
c	44	35.6	11.1	663	12	BG450476	BG450476 NF019E01D
c	45	35.6	11.1	691	12	BG451906	BG451906 NF101C08D

ALIGNMENTS

RESULT 1
BF942735
LOCUS BF942735 197 bp mRNA linear EST 15-SEP-2001
DEFINITION EST-CD15N-040 human CD15+ myeloid progenitor cells cDNA Library Homo sapiens cDNA 3', mRNA sequence.
ACCESSION BF942735
VERSION BF942735.1 GI:15624134
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 197)
AUTHORS Chen,J., Lee,S., Zhou,G., Rowley,J.D. and Wang,S.M.
TITLE A high-throughput GLGI procedure for converting large number of SAGE tag sequences into 3' ESTs
JOURNAL Unpublished (2001)
COMMENT Contact: Wang SM
Hem/Onc
University of Chicago Medical Center
5841 S. Maryland Ave., MC2115, Chicago, IL 60637, USA
Tel: 773-702-6788
Fax: 773-702-3002
Email: swang1@midway.uchicago.edu
This EST fragment was amplified from human CD15+ myeloid progenitor cells cDNA Library with GLGI technique (Generation of Longer cDNA fragments from SAGE tags for Gene Identification, Proc. Natl. Acad. Sci. USA 97, 349, 2000), which starts from the 3' end till the last CATG site of the target cDNA sequence.

```
           Seq primer: M13 Forward.
FEATURES             Location/Qualifiers
     source          1. .197
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="human CD15+ myeloid progenitor cells cDNA
                     Library"
                     /tissue_type="bone marrow"
                     /cell_type="CD15+ myeloid progenitor cells"
                     /note="Organ: thyroid; Vector: pAMP10; mRNA made from
                     thyroid carcinoma, cDNA made by oligo-dT  priming.
                     Non-directionally cloned into UDG sites. Size-selected on
                     agarose gel, average insert size 500 bp. Primary library.
                     cDNA Library Preparation: David B. Krizman, Ph.D.
                     REFERENCE: Krizman et al. (1996) Cancer Research
                     56:5380-5383."
BASE COUNT       44 a     38 c     37 g     78 t
ORIGIN

  Query Match             58.0%;  Score 186.2;  DB 12;  Length 197;
  Best Local Similarity   98.4%;  Pred. No. 4.3e-37;
  Matches  188;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      29 TTGTGTTTTCCTATATTACTATGTTCTGTTCCATTCAAAAAACCGCCTTGCAGACCACAG 88
           || | ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5 TTTTTTTTCCCTATATTACTATGTTCTGTTCCATTCAAAAAACCGCCTTGCAGACCACAG 64

Qy      89 AAGTAAGGAATTGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTTATAGTGT 148
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65 AAGTAAGGAATTGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTTATAGTGT 124

Qy     149 TCTCTGATGCCATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCTTCCGGG 208
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     125 TCTCTGATGCCATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCTTCCGGG 184

Qy     209 TGGAAATACCA 219
           ||||||||||
Db     185 TGGAAATACCA 195


RESULT 2
BQ228832
LOCUS       BQ228832                841 bp    mRNA    linear   EST 02-MAY-2002
DEFINITION  AGENCOURT_7522585 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6051251
            5', mRNA sequence.
ACCESSION   BQ228832
VERSION     BQ228832.1  GI:20410232
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 841)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technolcgies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13304  row: m  column: 12
            High quality sequence start: 87
            High quality sequence stop: 603.
FEATURES             Location/Qualifiers
     source          1. .841
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6051251"
                     /clone_lib="NIH_MGC_72"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
BASE COUNT      240 a    160 c    170 g    269 t      2 others
ORIGIN

  Query Match             49.5%;  Score 159;  DB 14;  Length 841;
  Best Local Similarity   71.2%;  Pred. No. 3.4e-30;
  Matches  210;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
           ||| | || |||   || ||| ||||||| || |||||||||   |  ||||| | |
Db     335 GGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGT 394

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
            ||||  |||  ||| |     | || ||| || |||||    ||  || |||| ||
Db     395 GTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATC 454

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
           |||| || ||||| || |||||||| ||  |||||||  | ||||||||| ||  | |||
Db     455 CTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTT 514

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
           ||||| ||| | ||||| ||  | | ||| |||||||||  || || || || ||| ||
Db     515 GTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTA 574

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCA 295
           ||||||||| | || ||  |||||||||||||||| ||||| ||||||||||| |
Db     575 GTGATTTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGA 629


RESULT 3
BG304121/c
LOCUS       BG304121                744 bp    mRNA    linear   EST 23-FEB-2001
DEFINITION  fl30f05.x1 Sugano Kawakami zebrafish DRA Danio rerio cDNA clone
            3815552 3' similar to TR:Q9VBP0 Q9VBP0 CG5042 PROTEIN. ;, mRNA
            sequence.
ACCESSION   BG304121
VERSION     BG304121.1  GI:13101648
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 744)
  AUTHORS   Sugano,S., Kawakami,K., Johnson,S., Li,F., Marra,M., Eddy,S.,
            Hillier,L., Clifton,S., Allen,M., Gibbons,M., Jost,S., Kucaba,T.,
            Martin,J., Pape,D., Steptoe,M., Underwood,K., Theising,B., Ritter
            ,E., Bowers,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: S.L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA
            Sequencing by: Washington University Genome Sequencing Center Clone
            distribution information can be found through the I.M.A.G.E.
            Consortium/LLNL, send email to: info@image.llnl.gov
            Seq primer: T7 from Gibco
            High quality sequence stop: 510.
FEATURES             Location/Qualifiers
     source          1. .744
                     /organism="Danio rerio"
                     /strain="AB"
                     /db_xref="taxon:7955"
```

/clone="3815552"
/clone_lib="Sugano Kawakami zebrafish DRA"
/sex="mixed (one male and one female, including unfertilized eggs)"
/dev_stage="adult"
/lab_host="DH10B (phage resistant)"
/note="Vector: pME18S-FL3; Site_1: DraIII (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCCTTTTTTTTTTTTTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTTGGCCTACTGG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science) and kindly donated by Dr. Koichi Kawakami. Custom primers for sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end primer CGACCTGCAGCTCGAGCACA. "
BASE COUNT 204 a 158 c 186 g 196 t
ORIGIN

Query Match 48.2%; Score 154.6; DB 12; Length 744;
Best Local Similarity 67.6%; Pred. No. 4.4e-29;
Matches 217; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

```
Qy      1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60
          ||  | | |  |||||||||| |||||||| ||| | || ||    |  ||||| |  |||
Db    738 GGACTTACCCGGCTGGCTTTCCTCATCATCGTGGTCTCGTACTCCAGCATGTTTTACTCC 679

Qy     61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
          ||  | || || | | |     ||||||| |   |  | |  |    ||||| ||||| 
Db    678 ATCTATAAGACGGGCATCAGAGCCACAGAGCTGCCGGGTCGACTGCACAGAGACGTGGCC 619

Qy    121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
          | ||| | || ||||| || || ||||| |||||||| ||||||||||||  ||||
Db    618 ATGGCACACCGCTTCTTCTTCATCGTGTTTTCTGATGCCTTCTGCTGGATTCCCATATTC 559

Qy    181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
          | || |||||||| || ||  |   |||||| |||||||  || || || || ||| |
Db    558 ATGGTCAAAATCCTCTCTCTAATGGAGGTGGAGATACCAGGAACCATCACATCTTGGGTG 499

Qy    241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300
          ||||| ||  |||| ||  | |||||||| || || || ||||| || || || || |||
Db    498 GTGATCTTCATCCTGCCCATCAACAGTGCCTTAAACCCCATCCTTTACACACTGACCACC 439

Qy    301 AACTTTTTTAAGGACAAGTTG 321
          |  ||||| |  ||  || ||
Db    438 AGTTTTTTCAGAGAGCAGGTG 418
```

RESULT 4 102(b)
AW436170
LOCUS AW436170 235 bp mRNA linear EST 09-JUL-2000
DEFINITION 75600 MARC 2PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION AW436170
VERSION AW436170.1 GI:6971476
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 235)
AUTHORS Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W. and Keele,J.W.
TITLE Design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 20 and -minmatch 12 options.
PCR PRimers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCCCAGTCACGACG
Plate: 34 row: F column: 22
Seq primer: ATTTAGGTGACACTATAG.
FEATURES Location/Qualifiers
source 1. .235
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2PIG"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI; Library made from pooled tissue from testis, ovary, endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT 49 a 64 c 45 g 77 t
ORIGIN

Query Match 47.4%; Score 152; DB 10; Length 235;
Best Local Similarity 87.0%; Pred. No. 2.1e-28;
Matches 167; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

```
Qy    130 CGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTTGTAGTTAAA 189
          | ||| |||||||||| |||||||||||||||||||||||||||||| ||||||||  |||||
Db      5 CTTTTTTTTTTTATAGGGTTCTCTGATGCCATCTGCTGGATTCCCGTATTTGTCATTAAA 64

Qy    190 ATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATAGTGATTTTT 249
          || || |||||||||||||||||||||||| | || | ||||||||   || ||||||
Db     65 ATTCTCTCCCTCTTCCGGGTGGAAATACCAGGCTCACTCACTTCCTGGGGGGTCATTTTT 124

Qy    250 TTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACCAACTTTTTT 309
          |||||||| || |||||||| ||||| |||||||| || ||||||||||||| ||||||
Db    125 TTCCTTCCGGTGAACAGTGCCTTGAACCCAATCCTTTACACTCTCACAACCAGCTTTTTC 184

Qy    310 AAGGACAAGTTG 321
          ||||||| ||||
Db    185 AAGGACACGTTG 196
```

RESULT 5
BF021857
LOCUS BF021857 456 bp mRNA linear EST 29-DEC-2000
DEFINITION uy58h08.y1 McCarrey Eddy round spermatid Mus musculus cDNA clone IMAGE:3663807 5' similar to TR:Q9VYG0 Q9VYG0 CG4187 PROTEIN. ;, mRNA sequence.
ACCESSION BF021857
VERSION BF021857.1 GI:10753189
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 456)
AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

```
            MGI:1424575
            Seq primer: Primer name ambiguous
            High quality sequence stop: 386.
FEATURES              Location/Qualifiers
     source           1. .456
                      /organism="Mus musculus"
                      /strain="CD-1"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:3663807"
                      /clone_lib="McCarrey Eddy round spermatid"
                      /sex="male"
                      /tissue_type="round spermatids, pooled from multiple mice"
                      /dev_stage="60 day"
                      /lab_host="DH10B (phage-resistant)"
                      /note="Organ: testis; Vector: pBluescript SK+ (Stratagene
                      ); Site_1: XhoII; Site_2: EcoRI; cDNA oligo dT-primed
                      [5'-(GA)10-ACTAGTCTCGAGTTTTTTTTTTTT-3'] and directionally
                      cloned  using 5' linkers 5'-AATTCGGCACGAG-3' and
                      5'-CTCGTGCCG-3'. Size selection of >400bp material gives
                      average insert size ranging from 1-2 kb. Library was mass
                      excised (from  lambda-UniZAP-XR) and resulting
                      single-stranded phagemids were prepped and tranformed
                      into DH10B. Library contains 98.5% recombinants.
                      References: J. Androl. 20:635-639 and Gene 25:263-269.
                      Library constructed and donated by J. McCarrey, Ph.D.
                      (Southwest Foundation for Biomedical Research, Dept. of
                      Genetics); excision done by E.M. Eddy, Ph.D. (National
                      Institutes of  Health, National Institute of Environmental
                      Health Sciences).  Original lambda-based library is
                      available through ATCC, catalog #63423."
BASE COUNT      120 a    110 c    111 g    115 t
ORIGIN

  Query Match             24.1%;  Score 77.4;  DB 12;  Length 456;
  Best Local Similarity   84.5%;  Pred. No. 1.7e-09;
  Matches   87;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    219 AGACACAATGACTTCCTGGATAGTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCC 278
          || |||||| |||||||||| ||| |||||||||||| || ||||| || || || ||
Db     36 AGGCACAATCACTTCCTGGATCGTGGTTTTTTTCCTTCCGGTGAACAGCGCCTTAAACCC 95

Qy    279 AATCCTCTATACTCTCACAACCAACTTTTTTAAGGACAAGTTG 321
          |||||||| ||||| || |||  ||||||||||||||||||
Db     96 CATCCTCTACACTCTGACGACCTCCTTTTTTAAGGACAAGTTG 138


RESULT 6
BB632910
LOCUS       BB632910                574 bp    mRNA    linear   EST 26-OCT-2001
DEFINITION  BB632910 RIKEN full-length enriched, adult male hypothalamus Mus
            musculus cDNA clone A230094D06 5', mRNA sequence.
ACCESSION   BB632910
VERSION     BB632910.1  GI:16469432
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 574)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
            , Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
            ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES              Location/Qualifiers
     source           1. .574
                      /organism="Mus musculus"
                      /db_xref="taxon:10090"
                      /clone="A230094D06"
                      /clone_lib="RIKEN full-length enriched, adult male
                      hypothalamus"
                      /sex="male"
                      /tissue_type="hypothalamus"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                      prepared and sequenced in Mouse Genome Encyclopedia
                      Project of Genome Exploration Research Group in Riken
                      Genomic Sciences Center and Genome Science Laboratory in
                      RIKEN. Division of Experimental Animal Research in Riken
                      contributed to prepare mouse tissues. 1st strand cDNA was
                      primed with a primer [5'
                      GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
                      prepared by using trehalose thermo-activated reverse
                      transcriptase and subsequently enriched for full-length by
                      cap-trapper. cDNA went through one round of normalization
                      to Rot = 20.0 and subtraction to Rot = 459.0. Second
                      strand cDNA was prepared with the primer adapter of
                      sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC
                      3']. cDNA was cleaved with XhoI and BamHI. Vector: a
                      modified pBluescript KS(+) after bulk excision from Lambda
                      FLC I. "
BASE COUNT      153 a    139 c    141 g    141 t
ORIGIN

  Query Match             24.1%;  Score 77.4;  DB 10;  Length 574;
  Best Local Similarity   84.5%;  Pred. No. 1.7e-09;
  Matches   87;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    219 AGACACAATGACTTCCTGGATAGTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCC 278
          || |||||| |||||||||| ||| |||||||||||| || ||||| || || || ||
Db     66 AGGCACAATCACTTCCTGGATCGTGGTTTTTTTCCTTCCGGTGAACAGCGCCTTAAACCC 125

Qy    279 AATCCTCTATACTCTCACAACCAACTTTTTTAAGGACAAGTTG 321
          |||||||| ||||| || |||  ||||||||||||||||||
Db    126 CATCCTCTACACTCTGACGACCTCCTTTTTTAAGGACAAGTTG 168
```

```
RESULT 7
CNS016W7/c
LOCUS       CNS016W7                1100 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC
            BACN16C04 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL107281
VERSION     AL107281.1  GI:5626586
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1100)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .1100
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACN16C04"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : T7"
BASE COUNT      208 a     92 c     73 g    194 t    533 others
ORIGIN

  Query Match             13.5%;  Score 43.2;  DB 17;  Length 1100;
  Best Local Similarity   26.5%;  Pred. No. 0.81;
  Matches   78;  Conservative   87;  Mismatches  129;  Indels    0;  Gaps    0;

Qy       3 TGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCCAT 62
Db    1053 TATSAWSTSWBWSSBVTBTSTWATSTMTTASTATBTTTSATVBWSATWTTHBASTNSVAT 994

Qy      63 TCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCTGT 122
Db     993 MVAAAAASAACABMTAVAAWTHASMATAYBASASAVYSSATVTSNCVSAAWCATAWMAMT 934

Qy     123 TGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTTGT 182
Db     933 TVTVTATSTCTTTTKTTTTHSTCTBTTTYTCWTTWCTBTCTSMMTTYHTWCATCTAYTYT 874

Qy     183 AGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATAGT 242
Db     873 ATYTTTTBMAAAKVTAYRTCTMYSTSAKYAAYAATABAWAABABAAWMHWAWAKKATTAT 814

Qy     243 GATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCAC 296
Db     813 GTSTAGWSTCKMGTSTBCTBAMMKAKCTKBKWBKCTAHKCACKBTMTSTVKMRM 760


RESULT 8
CNS04L3O
LOCUS       CNS04L3O                992 bp    DNA     linear   GSS 21-MAY-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
            117D12 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL295629
VERSION     AL295629.1  GI:8034209
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE   1  (bases 1 to 992)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 992)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
  TITLE     Charaterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 992)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1. .992
                     /organism="Tetraodon nigroviridis"
                     /db_xref="taxon:99883"
                     /clone="117D12"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0BG117DB06SP1-end :
                     PUC-Ori"
BASE COUNT      241 a    150 c    104 g    406 t     91 others
ORIGIN

  Query Match             12.6%;  Score 40.6;  DB 17;  Length 992;
  Best Local Similarity   38.1%;  Pred. No. 3.7;
  Matches  109;  Conservative   21;  Mismatches  156;  Indels    0;  Gaps    0;

Qy      22 CTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCCATTCAAAAAACCGCCTTGCAG 81
Db     305 CKTATCATAWTATWATTTKATTTCWCWTTCCCAAWTWCYTTCTACAACAGCTGCGCAACG 364

Qy      82 ACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTT 141
Db     365 TCKTTTCCAAATYTKCTATTTTCGGTAGATTTTCCCTTTTTTTTAATTTTTTTTTTTTTT 424

Qy     142 ATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTC 201
Db     425 TTTTTTTTTTATGTTTTTTGGTKTTCAGATTGWTTTWTCGTAKTTAKATACCATYTTTKC 484

Qy     202 TTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATAGTGATTTTTTTCCTTCCAGTT 261
Db     485 TTYCAGKTKCTTATTCCAACCKTTATTTTTCGCKKTTNNNTNTNNTTNNNTNNNNNNTT 544

Qy     262 AACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACCAACTTTT 307
Db     545 NNNNNTTWTTTWTCTCAGAAAATYCGTTWTTTCCCAATTTAATTWT 590


RESULT 9
BH446416
LOCUS       BH446416                660 bp    DNA     linear   GSS 12-DEC-2001
DEFINITION  BOHFU39TR BOHF Brassica oleracea genomic clone BOHFU39, DNA
            sequence.
ACCESSION   BH446416
VERSION     BH446416.1  GI:17632130
KEYWORDS    GSS.
SOURCE      Brassica oleracea.
```

```
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 660)
  AUTHORS   Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
  TITLE     Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL   Unpublished (2001)
COMMENT     Other_GSSs: BOHFU39TF
            Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .660
                     /organism="Brassica oleracea"
                     /strain="TO1000DH3"
                     /db_xref="taxon:3712"
                     /clone="BOHFU39"
                     /clone_lib="BOHF"
                     /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
                     genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT      175 a      95 c     137 g     253 t
ORIGIN

  Query Match             12.5%;  Score 40;  DB 17;  Length 660;
  Best Local Similarity   51.7%;  Pred. No. 5.3;
  Matches   91;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy      14 TGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCCATTCAAAAAACCG 73
           |||  |||  |||   | | |  ||| | ||| ||  ||| |||||||    ||  |  |
Db     349 TGGTGTTTGCCATATGTTTATGCTCCGAGATTGCTCGGTTATGTTCCAAAAGAATTATAG 408

Qy      74 CCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTT 133
           | |    |  |    || | |  | | || ||  ||      || || ||  ||
Db     409 TTTAGTTTGCTTCCTTTGTCATGTCCTTCTATTGATGAATGGAATTATTTTCATATTTTT 468

Qy     134 TCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTTGTAGTTAAA 189
           | |||||  || | || |||  | |    |  |  | || || | || | ||  ||
Db     469 TTTTTTTATTATTCTTTTCTTTTTCTTGGTATTTTACTCATGGAGTTCTTGTAGAA 524


RESULT 10
CNS00BEG/c
LOCUS       CNS00BEG                1101 bp    DNA     linear   GSS 04-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR23C02 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL056872
VERSION     AL056872.1  GI:4937540
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACR23C02"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
BASE COUNT      333 a     209 c     154 g     266 t     139 others
ORIGIN

  Query Match             12.2%;  Score 39.2;  DB 17;  Length 1101;
  Best Local Similarity   33.0%;  Pred. No. 8.4;
  Matches   70;  Conservative   47;  Mismatches   95;  Indels    0;  Gaps    0;

Qy      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120
           ||: : |  :  |  ::  ::   | | |  |:|  : || |:|| :  | :   ||: :
Db    1092 ATWTWCATKWGTGGKKKTAWKTWTAWAAATTTWATTWTTTTTDTTTDTKGGKGTATGKKK 1033

Qy     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180
           ::::  :  |: : :|||||||:|: |   :|  :|   | |  ||| ||  | : | |
Db    1032 KKKKKKTRGGGKGKKWTTTTTATRGKTTRKMYTTTWGTTTTTTWTTGGTTTTTTAKGTGT 973

Qy     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240
               |::  :  |||:   | :|  ||| |:  :     |  :  :| ||:|  | |:|
Db     972 AGTTTYVGTWATCTTKTTATGKTTKGGGGGKKKDGGKKGGGGMATWTAACKTRKTTGWTY 913

Qy     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTT 272
           :| :|||::||:::|   :    :||| |||
Db     912 CHGYYTTTKYTCYYYCTTTKSTGGRGTGRTTT 881


RESULT 11
CNS015ZE
LOCUS       CNS015ZE                1024 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC
            BACN15B13 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL106100
VERSION     AL106100.1  GI:5620062
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1024)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .1024
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACN15B13"
```

```
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : T7"
BASE COUNT      272 a    192 c    168 g    237 t    155 others
ORIGIN

  Query Match             12.1%;  Score 39;  DB 17;  Length 1024;
  Best Local Similarity   45.3%;  Pred. No. 9.4;
  Matches   48;  Conservative   23;  Mismatches   35;  Indels    0;  Gaps    0;

Qy     212 AAATACCAGACACAATGACTTCCTGGATAGTGATTTTTTTCCTTCCAGTTAACAGTGCTT 271
           :||||: | | | :| || | : :|  |  | ||||||:::::::|: ||     | ||
Db     822 MAATAMWAAAWAAWACGAATAMTWGTYTTYTCATTTTTYHHYHYYCHTTTYTTTWTTCTC 881

Qy     272 TGAATCCAATCCTCTATACTCTCACAACCAACTTTTTTAAGGACAA 317
           :|  : |:||| : ||: : ::|:||:||   ||||| || | ||
Db     882 NKATAMWAWTCCCVGATWTWWWMAMAABCATGGTTTTTSAGTAAAA 927


RESULT 12
AG148351
LOCUS       AG148351                681 bp    DNA     linear   GSS 09-JAN-2002
DEFINITION  Pan troglodytes DNA, clone: RP43-010I12.T7, genomic survey
            sequence.
ACCESSION   AG148351
VERSION     AG148351.1  GI:16678029
KEYWORDS    GSS.
SOURCE      Pan troglodytes male lymphocytes DNA, clone_lib:RPCI-43 Chimpanzee
            Male BAC Library clone:RP43-010I12.T7.
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library RPCI-43
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 681)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
            end was generated during the R&D process and may have higher chance
            of clone tracking errors.
             PRIMERS
                 Sequencing: T7
             LIBRARY
                 Vector    : pBACe3.6
                 R.Site 1  : EcoRI
                 R.Site 2  : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .681
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
                     /clone="RP43-010I12.T7"
                     /sex="male"
                     /cell_type="lymphocytes"
                     /clone_lib="RPCI-43 Chimpanzee Male BAC Library"
BASE COUNT      159 a    103 c     97 g    320 t      2 others
ORIGIN

  Query Match             11.9%;  Score 38.2;  DB 17;  Length 681;
  Best Local Similarity   52.1%;  Pred. No. 15;
  Matches   85;  Conservative    0;  Mismatches   78;  Indels    0;  Gaps    0;

Qy      99 TTGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGC 158
           || |||| |||||        | |||    ||   |   ||||||| | ||||||| | |
Db     107 TTATTTTTGAAGATGAAACTTTCTTGTTTCTCTGGTAGATTTTATAATTTTCTCTGTTAC 166

Qy     159 CATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACC 218
           |  ||  ||| ||   || |   |||| |  | |    | | |    ||||  ||
Db     167 TTAATTTTGTGTTCATGGTATTTTGAGTAAAGTTATGTTATTTTCCTTCTTGGACTTAAT 226

Qy     219 AGACACAATGACTTCCTGGATAGTGATTTTTTTCCTTCCAGTT 261
             |||  | | |     ||      ||||||| || ||||||
Db     227 TAGCACTGTCATTCTAAAGACTCATCTTTTTTTTCTCCCAGTT 269


RESULT 13
AL551397/c
LOCUS       AL551397                1073 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  AL551397 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI064YC04 5
            prime, mRNA sequence.
ACCESSION   AL551397
VERSION     AL551397.1  GI:12889305
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1073)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .1073
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DI064YC04"
                     /clone_lib="LTI_NFL006_PL2"
                     /tissue_type="placenta"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer.  Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      393 a    132 c    190 g    208 t    150 others
ORIGIN

  Query Match             11.8%;  Score 38;  DB 9;  Length 1073;
  Best Local Similarity   34.3%;  Pred. No. 17;
  Matches   68;  Conservative   40;  Mismatches   90;  Indels    0;  Gaps    0;

Qy      18 TTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCCATTCAAAAAACCGCCTT 77
           :||| |  |   |:||:::: |::::|  | |:||:|:||       :||    :   |
Db     363 WTTTTTTYTWCATKTGKWKKYTTWWWKTTTTTTKTTYTKTTTTKAAGGWAACCATKTKCT 304

Qy      78 GCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTT 137
            ::    ||: |  :     :|| |:||::     :  :   | :|: || |  |:  ||
Db     303 TYWTKWAACWAATTKTTTWKWTTYTKTTKKTKTTTWWTKCCTTTYTKGAAKTTTTKCATT 244

Qy     138 TTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTC 197
           ||:  : :| || | :  |    | | ||  : | ||:| ||  |  ||: | : ||||
Db     243 TTKCCKTKTTTTTTKKTTTTTACTTTCCTCCWGTKCTKTTTTACTTCTTWCATKWCTTTK 184

Qy     198 CCTCTTCCGGGTGGAAAT 215
           : | |||    :: || |
Db     183 YTTTTTCATCCWKTAATT 166


RESULT 14
FR0020632/c
```

```
LOCUS       FR0020632               598 bp    DNA     linear   GSS 10-DEC-1997
DEFINITION  F.rubripes GSS sequence, clone 042J05aA7, genomic survey sequence.
ACCESSION   AL013515
VERSION     AL013515.1  GI:2679883
KEYWORDS    GSS; genome survey sequence.
SOURCE      Takifugu rubripes.
  ORGANISM  Takifugu rubripes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Takifugu.
REFERENCE   1  (bases 1 to 598)
  AUTHORS   Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrania,Y.,
            Williams,G. and Brenner,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-DEC-1997) MRC Human Genome Mapping Project Resource
            Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgmp.mrc.ac.uk
COMMENT     Vector: pBluescript II KS
            V_type: phagemid
            PRIMER: KS
            DESCR:
            One pass dye-terminator sequencing of cosmid cloned genomic
            sequence.
FEATURES             Location/Qualifiers
     source          1. .598
                     /organism="Takifugu rubripes"
                     /db_xref="taxon:31033"
                     /clone="042J05aA7"
                     /clone_lib="cosmid 042J05"
BASE COUNT      182 a    114 c    127 g    165 t     10 others
ORIGIN
```

```
  Query Match             11.8%;  Score 37.8;  DB 17;  Length 598;
  Best Local Similarity   53.4%;  Pred. No. 19;
  Matches   78;  Conservative    0;  Mismatches   68;  Indels    0;  Gaps    0;

Qy     163 TGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGAC 222
           ||||| ||| || ||  | | |  |  ||  ||  |      || || |||||||||
Db     350 TGCTGAATTTTTGAATGCGCATTACATGTCACTTTACATCAAGGGTTGCAAATACCAGTT 291

Qy     223 ACAATGACTTCCTGGATAGTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATC 282
           | |||         |||  |   || |||||    |   ||| ||||  |   |   | |
Db     290 AAAATTGTGATGCAGATCATTTATTATTTCCCAAAAACCAACCGTGCGATTNTTATTAGC 231

Qy     283 CTCTATACTCTCACAACCAACTTTTT 308
           ||| | ||  | || || | |||||
Db     230 ATCTTTTCTTCCCCACCCCATTTTTT 205
```

```
RESULT 15
CNS016CF
LOCUS       CNS016CF                1201 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC
            BACN15I18 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL106569
VERSION     AL106569.1  GI:5622587
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .1201
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACN15I18"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : T7"
BASE COUNT      235 a    265 c    159 g    325 t    217 others
ORIGIN
```

```
  Query Match             11.8%;  Score 37.8;  DB 17;  Length 1201;
  Best Local Similarity   34.5%;  Pred. No. 19;
  Matches   51;  Conservative   39;  Mismatches   58;  Indels    0;  Gaps    0;

Qy     173 CTGTATTTGTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTT 232
           || | |||: |:|:   | : : : |:|: |  |:   | ||| : | : : :  | : |
Db     835 CTTTBTTTKGABTYTTCASYTYGYKCSTBSTKKGBNGTGCAATTSKATWTRRRGGGGSST 894

Qy     233 CCTGGATAGTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTC 292
           :  ::   |     |:::::|:|  |:: |  ||:||| :||  :|| ||:| : :|:
Db     895 TWKTTDKTTAGNGNAATWHHHYTYCTTTYHBCTTTGSTTTTWATATHATACTYTCYTYTY 954

Qy     293 TCACAACCAACTTTTTTAAGGACAAGTT 320
           | | || ::  ::|| | :|: ||| :|
Db     955 TAAAAAWYMCTWYTTATTWGSCCAAAYT 982
```

```
Search completed: May 26, 2003, 20:03:48
Job time : 1753 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 22, 2003, 12:33:39 ; Search time 74 Seconds
(without alignments)
192.673 Million cell updates/sec

Title: US-09-930-312-2
Perfect score: 549
Sequence: 1 GVNLLAFLIIVFSYITMFCS.....NSALNPILYTLTTNFFKDKL 107

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	549	100.0	107	23	AAE20148	Human novel G-prot
2	549	100.0	108	23	AAE20167	Human nGPCR-1079 v
3	549	100.0	109	23	AAE20168	Human nGPCR-1079 v
4	549	100.0	355	22	AAU04370	Human G-protein co
5	549	100.0	355	23	ABB06253	Human G protein-co
6	549	100.0	491	23	ABB06252	Human G protein-co
7	549	100.0	515	23	ABB06251	Human G protein-co
8	549	100.0	610	23	ABB06250	Human G protein-co
9	549	100.0	713	23	AAU97159	Human G-protein co
10	549	100.0	730	23	ABB06255	Human G protein-co
11	549	100.0	737	23	AAU97158	Human G-protein co
12	549	100.0	754	23	ABB06254	Human G protein-co
13	376	68.5	176	22	AAU87545	Novel central nerv
14	376	68.5	176	22	AAM99972	Human expressed po
15	376	68.5	176	22	ABB04062	Human musculoskele
16	376	68.5	176	22	ABB10249	Human cDNA SEQ ID
17	376	68.5	176	22	AAU18108	Novel human uterin
18	376	68.5	176	22	AAU18344	Human endocrine po
19	376	68.5	176	22	AAU18666	Renal and cardiova
20	376	68.5	176	22	AAU21646	Novel human neopla
21	376	68.5	188	22	AAM99945	Human expressed po
22	376	68.5	209	22	AAU87243	Novel central nerv
23	376	68.5	396	21	AAY57286	Human GPCR protein
24	376	68.5	396	22	AAE02498	Human CON222 G pro
25	376	68.5	722	20	AAY42171	Human LGR7 short f
26	376	68.5	757	20	AAY42170	Human LGR7 long fo
27	371	67.6	327	22	AAG80961	Human nGPCR57. Ho
28	370	67.4	140	21	AAB41526	Human ORFX ORF1290
29	349	63.6	157	22	AAB35407	Human HGRL101 G-pr
30	234	42.6	334	22	ABB60463	Drosophila melanog
31	234	42.6	334	22	AAU38929	Drosophila G-prote
32	225	41.0	359	22	ABB61216	Drosophila melanog
33	225	41.0	359	22	AAU38931	Drosophila G-prote
34	210	38.3	1115	23	AAM47959	Lymnaea stagnalis
35	142.5	26.0	685	12	AAR11331	Human luteinising
36	142	25.9	861	22	AAB86961	D. melanogaster pe
37	142	25.9	1300	22	ABB64083	Drosophila melanog
38	138.5	25.2	320	16	AAR79450	Rat A3 adenosine r
39	136.5	24.9	361	14	AAR30496	N-terminal of LH r
40	136.5	24.9	386	14	AAR30500	N-terminal of LH r
41	136.5	24.9	638	14	AAR30499	N-terminal of LH r
42	136.5	24.9	692	14	AAR30503	N-terminal of LH r
43	136.5	24.9	693	14	AAR30510	N-terminal of LH r
44	136.5	24.9	698	14	AAR30505	N-terminal of LH r
45	136.5	24.9	699	14	AAR30517	N-terminal of LH r

ALIGNMENTS

RESULT 1
AAE20148
ID AAE20148 standard; Protein; 107 AA.
XX
AC AAE20148;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human novel G-protein coupled receptor (nGPCR)-1079 protein.
XX
KW Human; novel G-protein coupled receptor; nGPCR-1079; viral infection;
KW gene; therapy; human immunodeficiency virus; HIV; pain; migraine;
KW central nervous system disorder; stroke; manic depression; obesity;
KW metabolic disorder; anorexia; cardiovascular disorder; type 2 diabetes;
KW cancer; myocardial infarction; hypotension; degenerative disorder;
KW Parkinson's disease; Alzheimer's disease; neurological disorder;
KW schizophrenia; anxiety; inflammatory condition; rheumatoid arthritis;
KW thyroid disorder; autoimmune disorder; hormonal disorder; renal failure;
KW psoriasis; movement disorder; analgesic; cytostatic; neuroprotective;
KW anorectic; hypotensive; hypertensive; tranquilliser; anticonvulsant;
KW metabolic; neuroleptic; thrombolytic; cardiant; immunosuppressive.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 4..26
FT /label= Transmembrane_domain
FT Domain 45..68
FT /label= Transmembrane_domain
FT Domain 77..99
FT /label= Transmembrane_domain
XX
PN WO200214496-A2.

XX
PD 21-FEB-2002.
XX
PF 15-AUG-2001; 2001WO-US25513.
XX
PR 15-AUG-2000; 2000US-225262P.
XX
PA (PHAA) PHARMACIA & UPJOHN CO.
XX
PI Lind P;
XX
DR WPI; 2002-269192/31.
XX
PT An isolated nucleic acid molecule encoding novel G-protein coupled
PT receptor polypeptide which is useful for treating obesity, diabetes,
PT Parkinson's disease, manic depression, migraine, rheumatoid arthritis
PT -
XX
PS Claim 31; Page 60; 93pp; English.
XX
CC The patent discloses novel G-protein coupled receptor (nGPCR)-1079 genes
CC and their corresponding proteins. Sequences of the invention are useful
CC for treating diseases such as viral infections caused by human immuno-
CC deficiency virus (HIV)-1 or 2, central nervous system (CNS) disorders
CC (e.g. pain, including migraine, stroke, manic depression), metabolic
CC disorders (e.g. obesity and anorexia), cancers, cardiovascular disorders
CC (e.g. type 2 diabetes, myocardial infarction, hypotension), degenerative
CC disorders (e.g. Parkinson's disease, Alzheimer's disease), neurological
CC disorders (e.g. schizophrenia and anxiety), inflammatory conditions,
CC rheumatoid arthritis, thyroid disorders, autoimmune disorders, hormonal
CC disorders, renal failure, psoriasis and movement disorders. The present
CC sequence is human nGPCR-1079 protein.
XX
SQ Sequence 107 AA;

Query Match 100.0%; Score 549; DB 23; Length 107;
Best Local Similarity 100.0%; Pred. No. 3.5e-66;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60

Qy 61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
|||||||||||||||||||||||||||||||||||||||||||||||
Db 61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107

RESULT 2
AAE20167
ID AAE20167 standard; Protein; 108 AA.
XX
AC AAE20167;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human nGPCR-1079 variant protein, Met-nGPCR-1079.
XX
KW Human; novel G-protein coupled receptor; nGPCR-1079; viral infection;
KW gene; therapy; human immunodeficiency virus; HIV; pain; migraine;
KW central nervous system disorder; stroke; manic depression; obesity;
KW metabolic disorder; anorexia; cardiovascular disorder; type 2 diabetes;
KW cancer; myocardial infarction; hypotension; degenerative disorder;
KW Parkinson's disease; Alzheimer's disease; neurological disorder;
KW schizophrenia; anxiety; inflammatory condition; rheumatoid arthritis;
KW thyroid disorder; autoimmune disorder; hormonal disorder; renal failure;
KW psoriasis; movement disorder; analgesic; cytostatic; neuroprotective;
KW anorectic; hypotensive; hypertensive; tranquilliser; anticonvulsant;
KW metabolic; neuroleptic; thrombolytic; cardiant; immunosuppressive;
KW mutant; mutein; variant.
XX
OS Homo sapiens.
OS Synthetic.

XX
PN WO200214496-A2.
XX
PD 21-FEB-2002.
XX
PF 15-AUG-2001; 2001WO-US25513.
XX
PR 15-AUG-2000; 2000US-225262P.
XX
PA (PHAA) PHARMACIA & UPJOHN CO.
XX
PI Lind P;
XX
DR WPI; 2002-269192/31.
XX
PT An isolated nucleic acid molecule encoding novel G-protein coupled
PT receptor polypeptide which is useful for treating obesity, diabetes,
PT Parkinson's disease, manic depression, migraine, rheumatoid arthritis
PT -
XX
PS Disclosure; Page -; 93pp; English.
XX
CC The patent discloses novel G-protein coupled receptor (nGPCR)-1079 genes
CC and their corresponding proteins. Sequences of the invention are useful
CC for treating diseases such as viral infections caused by human immuno-
CC deficiency virus (HIV)-1 or 2, central nervous system (CNS) disorders
CC (e.g. pain, including migraine, stroke, manic depression), metabolic
CC disorders (e.g. obesity and anorexia), cancers, cardiovascular disorders
CC (e.g. type 2 diabetes, myocardial infarction, hypotension), degenerative
CC disorders (e.g. Parkinson's disease, Alzheimer's disease), neurological
CC disorders (e.g. schizophrenia and anxiety), inflammatory conditions,
CC rheumatoid arthritis, thyroid disorders, autoimmune disorders, hormonal
CC disorders, renal failure, psoriasis and movement disorders. The present
CC sequence is human nGPCR-1079 variant protein, Met-nGPCR-1079. This
CC sequence is generated by the addition of a Met residue at position -1
CC of human wild-type nGPCR-1079 protein.
CC Note: This sequence is not shown in the specification but is derived from
CC human wild-type nGPCR-1079 protein shown as SEQ ID NO: 2 in page 60 of
CC the specification (AAE20148).
XX
SQ Sequence 108 AA;

Query Match 100.0%; Score 549; DB 23; Length 108;
Best Local Similarity 100.0%; Pred. No. 3.5e-66;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 61

Qy 61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
|||||||||||||||||||||||||||||||||||||||||||||||
Db 62 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 108

RESULT 3
AAE20168
ID AAE20168 standard; Protein; 109 AA.
XX
AC AAE20168;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human nGPCR-1079 variant protein, Met-Lys-nGPCR-1079.
XX
KW Human; novel G-protein coupled receptor; nGPCR-1079; viral infection;
KW gene; therapy; human immunodeficiency virus; HIV; pain; migraine;
KW central nervous system disorder; stroke; manic depression; obesity;
KW metabolic disorder; anorexia; cardiovascular disorder; type 2 diabetes;
KW cancer; myocardial infarction; hypotension; degenerative disorder;
KW Parkinson's disease; Alzheimer's disease; neurological disorder;
KW schizophrenia; anxiety; inflammatory condition; rheumatoid arthritis;
KW thyroid disorder; autoimmune disorder; hormonal disorder; renal failure;

KW psoriasis; movement disorder; analgesic; cytostatic; neuroprotective;
KW anorectic; hypotensive; hypertensive; tranquilliser; anticonvulsant;
KW metabolic; neuroleptic; thrombolytic; cardiant; immunosuppressive;
KW mutant; mutein; variant.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200214496-A2.
XX
PD 21-FEB-2002.
XX
PF 15-AUG-2001; 2001WO-US25513.
XX
PR 15-AUG-2000; 2000US-225262P.
XX
PA (PHAA) PHARMACIA & UPJOHN CO.
XX
PI Lind P;
XX
DR WPI; 2002-269192/31.
XX
PT An isolated nucleic acid molecule encoding novel G-protein coupled
PT receptor polypeptide which is useful for treating obesity, diabetes,
PT Parkinson's disease, manic depression, migraine, rheumatoid arthritis
PT -
XX
PS Disclosure; Page -; 93pp; English.
XX
CC The patent discloses novel G-protein coupled receptor (nGPCR)-1079 genes
CC and their corresponding proteins. Sequences of the invention are useful
CC for treating diseases such as viral infections caused by human immuno-
CC deficiency virus (HIV)-1 or 2, central nervous system (CNS) disorders
CC (e.g. pain, including migraine, stroke, manic depression), metabolic
CC disorders (e.g. obesity and anorexia), cancers, cardiovascular disorders
CC (e.g. type 2 diabetes, myocardial infarction, hypotension), degenerative
CC disorders (e.g. Parkinson's disease, Alzheimer's disease), neurological
CC disorders (e.g. schizophrenia and anxiety), inflammatory conditions,
CC rheumatoid arthritis, thyroid disorders, autoimmune disorders, hormonal
CC disorders, renal failure, psoriasis and movement disorders. The present
CC sequence is human nGPCR-1079 variant protein, Met-Lys-nGPCR-1079. This
CC sequence is generated by the addition of Met-Lys residues at the amino
CC terminus (positions -2 and -1) of human wild-type nGPCR-1079 protein.
CC Note: This sequence is not shown in the specification but is derived from
CC human wild-type nGPCR-1079 protein shown as SEQ ID NO: 2 in page 60 of
CC the specification (AAE20148).
XX
SQ Sequence 109 AA;

Query Match 100.0%; Score 549; DB 23; Length 109;
Best Local Similarity 100.0%; Pred. No. 3.6e-66;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
Db 3 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 62

Qy 61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
Db 63 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 109

RESULT 4
AAU04370
ID AAU04370 standard; Protein; 355 AA.
XX
AC AAU04370;
XX
DT 23-OCT-2001 (first entry)
XX
DE Human G-protein coupled receptor, hRUP16.
XX
KW Human; G-protein coupled receptor; GPCR; hRUP16; agonist;
KW inverse agonist; lung cancer.
XX
OS Homo sapiens.
XX
PN WO200136471-A2.
XX
PD 25-MAY-2001.
XX
PF 16-NOV-2000; 2000WO-US31509.
XX
PR 17-NOV-1999; 99US-0166088.
PR 17-NOV-1999; 99US-0166099.
PR 17-NOV-1999; 99US-0166369.
PR 23-DEC-1999; 99US-0171900.
PR 23-DEC-1999; 99US-0171901.
PR 23-DEC-1999; 99US-0171902.
PR 11-FEB-2000; 2000US-0181749.
PR 14-MAR-2000; 2000US-0189258.
PR 14-MAR-2000; 2000US-0189259.
PR 10-APR-2000; 2000US-0195898.
PR 10-APR-2000; 2000US-0195899.
PR 10-APR-2000; 2000US-0196078.
PR 28-APR-2000; 2000US-0200419.
PR 12-MAY-2000; 2000US-0203630.
PR 12-JUN-2000; 2000US-0210741.
PR 12-JUN-2000; 2000US-0210982.
PR 21-AUG-2000; 2000US-0226760.
PR 26-SEP-2000; 2000US-0235418.
PR 26-SEP-2000; 2000US-0235779.
PR 20-OCT-2000; 2000US-0242332.
PR 20-OCT-2000; 2000US-0242343.
XX
PA (AREN-) ARENA PHARM INC.
XX
PI Chen R, Dang HT, Lowitz KP;
XX
DR WPI; 2001-355616/37.
DR N-PSDB; AAS07943.
XX
PT Endogenous and non-endogenous versions of human G-protein coupled
PT receptors for direct identification of candidate compounds as agonists,
PT inverse agonists or partial agonists for use as therapeutic agents -
XX
PS Claim 33; Page 105-106; 160pp; English.
XX
CC The sequence represents a human G-protein coupled receptor (GPCR),
CC hRUP16. The endogenous and non-endogenous, constitutively activated
CC versions of human G-protein coupled receptors (GPCR), are useful for
CC direct identification of candidate compounds as receptor agonists,
CC inverse agonists or partial agonists having applicability as therapeutic
CC agents for treating diseases related to GPCR, e.g. lung cancer.
CC Non-endogenous version of human GPCRs are also utilized in research
CC settings and in vitro and in vivo system, incorporating GPCRs can be
CC utilised to elucidate and understand the roles these receptors
CC play in the human condition, both normal and diseased.
XX
SQ Sequence 355 AA;

Query Match 100.0%; Score 549; DB 22; Length 355;
Best Local Similarity 100.0%; Pred. No. 1.6e-65;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
Db 197 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 256

Qy 61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
Db 257 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 303

RESULT 5
ABB06253

```
ID   ABB06253 standard; Protein; 355 AA.
XX
AC   ABB06253;
XX
DT   23-MAY-2002  (first entry)
XX
DE   Human G protein-coupled receptor TGR17-4 protein SEQ ID NO:10.
XX
KW   Human; G protein-coupled receptor; TGR17-4; nootropic; antiinflammatory;
KW   vasotropic; immunomodulator; cytostatic; gene therapy; protein therapy;
KW   neurological; inflammatory; circulatory; degenerative; immune system;
KW   digestive disease; cancer.
XX
OS   Homo sapiens.
XX
PN   WO200204640-A1.
XX
PD   17-JAN-2002.
XX
PF   06-JUL-2001; 2001WO-JP05878.
XX
PR   07-JUL-2000; 2000JP-0211989.
PR   18-DEC-2000; 2000JP-0383794.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
PI   Moriya T,  Ito T,  Shintani Y,  Miyajima N;
XX
DR   WPI; 2002-179706/23.
DR   N-PSDB; ABL40194.
XX
PT   G-protein coupled receptor protein TGR17 of human origin and DNA
PT   encoding it for diagnosis and treatment of cancer and circulatory and
PT   other diseases associated with its expression -
XX
PS   Claim 1; Fig 8; 145pp; Japanese.
XX
CC   The present invention describes a human guanine nucleotide binding
CC   protein (G protein) coupled receptor protein designated TGR17, which has
CC   5 variant forms (TGR17-1 to TGR17-5) of the receptor protein having
CC   additional residues at the N-terminal. The TGR17 proteins have nootropic,
CC   antiinflammatory, vasotropic, immunomodulator and cytostatic activities.
CC   The TGR17 polynucleotides and protein can be used in gene therapy and
CC   protein therapy. G-protein coupled receptor proteins are cell membrane
CC   proteins mediating the cellular response to a large variety of signalling
CC   molecules. The TGR17 polynucleotides and proteins can be used in the
CC   diagnosis, treatment and prevention of diseases including neurological,
CC   inflammatory, circulatory, degenerative, immune system and digestive
CC   diseases and cancer. The present sequence represents human TGR17-4 from
CC   the present invention.
XX
SQ   Sequence   355 AA;

  Query Match             100.0%;  Score 549;  DB 23;  Length 355;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-65;
  Matches  107;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     197 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 256

Qy      61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
           |||||||||||||||||||||||||||||||||||||||||||||||
Db     257 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 303


RESULT 6
ABB06252
ID   ABB06252 standard; Protein; 491 AA.
XX
AC   ABB06252;
XX
DT   23-MAY-2002  (first entry)
XX
DE   Human G protein-coupled receptor TGR17-3 protein SEQ ID NO:7.
XX
KW   Human; G protein-coupled receptor; TGR17-3; nootropic; antiinflammatory;
KW   vasotropic; immunomodulator; cytostatic; gene therapy; protein therapy;
KW   neurological; inflammatory; circulatory; degenerative; immune system;
KW   digestive disease; cancer.
XX
OS   Homo sapiens.
XX
PN   WO200204640-A1.
XX
PD   17-JAN-2002.
XX
PF   06-JUL-2001; 2001WO-JP05878.
XX
PR   07-JUL-2000; 2000JP-0211989.
PR   18-DEC-2000; 2000JP-0383794.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
PI   Moriya T,  Ito T,  Shintani Y,  Miyajima N;
XX
DR   WPI; 2002-179706/23.
DR   N-PSDB; ABL40192.
XX
PT   G-protein coupled receptor protein TGR17 of human origin and DNA
PT   encoding it for diagnosis and treatment of cancer and circulatory and
PT   other diseases associated with its expression -
XX
PS   Claim 3; Fig 7; 145pp; Japanese.
XX
CC   The present invention describes a human guanine nucleotide binding
CC   protein (G protein) coupled receptor protein designated TGR17, which has
CC   5 variant forms (TGR17-1 to TGR17-5) of the receptor protein having
CC   additional residues at the N-terminal. The TGR17 proteins have nootropic,
CC   antiinflammatory, vasotropic, immunomodulator and cytostatic activities.
CC   The TGR17 polynucleotides and protein can be used in gene therapy and
CC   protein therapy. G-protein coupled receptor proteins are cell membrane
CC   proteins mediating the cellular response to a large variety of signalling
CC   molecules. The TGR17 polynucleotides and proteins can be used in the
CC   diagnosis, treatment and prevention of diseases including neurological,
CC   inflammatory, circulatory, degenerative, immune system and digestive
CC   diseases and cancer. The present sequence represents human TGR17-3 from
CC   the present invention.
XX
SQ   Sequence   491 AA;

  Query Match             100.0%;  Score 549;  DB 23;  Length 491;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-65;
  Matches  107;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     333 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 392

Qy      61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
           |||||||||||||||||||||||||||||||||||||||||||||||
Db     393 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 439


RESULT 7
ABB06251
ID   ABB06251 standard; Protein; 515 AA.
XX
AC   ABB06251;
XX
DT   23-MAY-2002  (first entry)
XX
DE   Human G protein-coupled receptor TGR17-2 protein SEQ ID NO:5.
XX
KW   Human; G protein-coupled receptor; TGR17-2; nootropic; antiinflammatory;
KW   vasotropic; immunomodulator; cytostatic; gene therapy; protein therapy;
```

```
KW   neurological; inflammatory; circulatory; degenerative; immune system;
KW   digestive disease; cancer.
XX
OS   Homo sapiens.
XX
PN   WO200204640-A1.
XX
PD   17-JAN-2002.
XX
PF   06-JUL-2001; 2001WO-JP05878.
XX
PR   07-JUL-2000; 2000JP-0211989.
PR   18-DEC-2000; 2000JP-0383794.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
PI   Moriya T,  Ito T,  Shintani Y,  Miyajima N;
XX
DR   WPI; 2002-179706/23.
DR   N-PSDB; ABL40191.
XX
PT   G-protein coupled receptor protein TGR17 of human origin and DNA
PT   encoding it for diagnosis and treatment of cancer and circulatory and
PT   other diseases associated with its expression -
XX
PS   Claim 3; Fig 6; 145pp; Japanese.
XX
CC   The present invention describes a human guanine nucleotide binding
CC   protein (G protein) coupled receptor protein designated TGR17, which has
CC   5 variant forms (TGR17-1 to TGR17-5) of the receptor protein having
CC   additional residues at the N-terminal. The TGR17 proteins have nootropic,
CC   antiinflammatory, vasotropic, immunomodulator and cytostatic activities.
CC   The TGR17 polynucleotides and protein can be used in gene therapy and
CC   protein therapy. G-protein coupled receptor proteins are cell membrane
CC   proteins mediating the cellular response to a large variety of signalling
CC   molecules. The TGR17 polynucleotides and proteins can be used in the
CC   diagnosis, treatment and prevention of diseases including neurological,
CC   inflammatory, circulatory, degenerative, immune system and digestive
CC   diseases and cancer. The present sequence represents human TGR17-2 from
CC   the present invention.
XX
SQ   Sequence   515 AA;

  Query Match             100.0%;  Score 549;  DB 23;  Length 515;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-65;
  Matches  107;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     357 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 416

Qy      61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
           |||||||||||||||||||||||||||||||||||||||||||||||
Db     417 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 463


RESULT 8
ABB06250
ID   ABB06250 standard; Protein; 610 AA.
XX
AC   ABB06250;
XX
DT   23-MAY-2002  (first entry)
XX
DE   Human G protein-coupled receptor TGR17-1 protein SEQ ID NO:1.
XX
KW   Human; G protein-coupled receptor; TGR17-1; nootropic; antiinflammatory;
KW   vasotropic; immunomodulator; cytostatic; gene therapy; protein therapy;
KW   neurological; inflammatory; circulatory; degenerative; immune system;
KW   digestive disease; cancer.
XX
OS   Homo sapiens.
XX
PN   WO200204640-A1.
XX
PD   17-JAN-2002.
XX
PF   06-JUL-2001; 2001WO-JP05878.
XX
PR   07-JUL-2000; 2000JP-0211989.
PR   18-DEC-2000; 2000JP-0383794.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
PI   Moriya T,  Ito T,  Shintani Y,  Miyajima N;
XX
DR   WPI; 2002-179706/23.
DR   N-PSDB; ABL40188.
XX
PT   G-protein coupled receptor protein TGR17 of human origin and DNA
PT   encoding it for diagnosis and treatment of cancer and circulatory and
PT   other diseases associated with its expression -
XX
PS   Claim 3; Fig 5; 145pp; Japanese.
XX
CC   The present invention describes a human guanine nucleotide binding
CC   protein (G protein) coupled receptor protein designated TGR17, which has
CC   5 variant forms (TGR17-1 to TGR17-5) of the receptor protein having
CC   additional residues at the N-terminal. The TGR17 proteins have nootropic,
CC   antiinflammatory, vasotropic, immunomodulator and cytostatic activities.
CC   The TGR17 polynucleotides and protein can be used in gene therapy and
CC   protein therapy. G-protein coupled receptor proteins are cell membrane
CC   proteins mediating the cellular response to a large variety of signalling
CC   molecules. The TGR17 polynucleotides and proteins can be used in the
CC   diagnosis, treatment and prevention of diseases including neurological,
CC   inflammatory, circulatory, degenerative, immune system and digestive
CC   diseases and cancer. The present sequence represents human TGR17-1 from
CC   the present invention.
XX
SQ   Sequence   610 AA;

  Query Match             100.0%;  Score 549;  DB 23;  Length 610;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-65;
  Matches  107;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     452 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 511

Qy      61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
           |||||||||||||||||||||||||||||||||||||||||||||||
Db     512 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 558


RESULT 9
AAU97159
ID   AAU97159 standard; Protein; 713 AA.
XX
AC   AAU97159;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human G-protein coupled receptor HGPRBMY5 splice variant.
XX
KW   Human; G-protein coupled receptor; GPCR; HGPRBMY5; colon; brain;
KW   ovary; thymus; lung; immune system; cancer; immune disorder;
KW   neurological disorder; infection; human immunodeficiency virus; HIV;
KW   antiallergic; antiasthmatic; dermatological; antiarteriosclerotic;
KW   antiulcer; antidiabetic; nephrotropic; osteopathic; antiarthritic;
KW   antiinflammatory; antirheumatic; antithyroid; cytostatic; vulnerary;
KW   virucide; antibacterial; antifungal; antiparasitic; protozoacide;
KW   antihelminthic; nootropic; neuroprotective; antidepressant;
KW   anticonvulsant; antiparkinsonian; neuroleptic; anti-HIV; receptor.
XX
OS   Homo sapiens.
XX
```

PN WO200226824-A2.
XX
PD 04-APR-2002.
XX
PF 26-SEP-2001; 2001WO-US30365.
XX
PR 27-SEP-2000; 2000US-235713P.
PR 16-JAN-2001; 2001US-261781P.
PR 19-JUL-2001; 2001US-306605P.
PR 03-AUG-2001; 2001US-310436P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Feder JN, Mintier G, Ramanathan CS, Hawken DR;
XX
DR WPI; 2002-435196/46.
DR N-PSDB; ABK51947.
XX
PT Novel G protein-coupled receptor, HGPRBMY5 polypeptide, useful for
PT treatment of AIDS, allergies, asthma, atherosclerosis, ulcerative
PT colitis, atopic dermatitis, diabetes mellitus, glomerulonephritis,
PT osteoarthritis -
XX
PS Claim 11; Fig 6; 148pp; English.
XX
CC The present invention relates to the isolation of a novel human
CC G-protein coupled receptor (GPCR) (HGPRBMY5), and the polynucleotide
CC sequence encoding it. The HGPRBMY5 polypeptide and polynucleotide
CC are useful for preventing, treating or ameliorating a disease,
CC disorder or condition related to the colon, brain, ovaries, thymus,
CC lungs or imm immune system. They are particularly useful for the
CC treatment or prevention of cancers, immune disorders, neurological
CC disorders, and diseases related to the brain, ovaries, thymus or
CC lungs. The polynucleotide sequence is useful for diagnosing or
CC determining susceptibility to infections such as bacterial, fungal,
CC protozoan and viral infections, particularly infections caused by
CC human immunodeficiency virus (HIV or HIV-2). The present sequence
CC represents a splice variant of human GPCR HGPRBMY5.
XX
SQ Sequence 713 AA;

Query Match 100.0%; Score 549; DB 23; Length 713;
Best Local Similarity 100.0%; Pred. No. 4e-65;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 555 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 614

Qy 61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
|||||||||||||||||||||||||||||||||||||||||||||
Db 615 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 661

RESULT 10
ABB06255
ID ABB06255 standard; Protein; 730 AA.
XX
AC ABB06255;
XX
DT 23-MAY-2002 (first entry)
XX
DE Human G protein-coupled receptor TGR17-6 protein SEQ ID NO:15.
XX
KW Human; G protein-coupled receptor; TGR17-6; nootropic; antiinflammatory;
KW vasotropic; immunomodulator; cytostatic; gene therapy; protein therapy;
KW neurological; inflammatory; circulatory; degenerative; immune system;
KW digestive disease; cancer.
XX
OS Homo sapiens.
XX
PN WO200204640-A1.
XX
PD 17-JAN-2002.
XX
PF 06-JUL-2001; 2001WO-JP05878.
XX
PR 07-JUL-2000; 2000JP-0211989.
PR 18-DEC-2000; 2000JP-0383794.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Moriya T, Ito T, Shintani Y, Miyajima N;
XX
DR WPI; 2002-179706/23.
DR N-PSDB; ABL40197.
XX
PT G-protein coupled receptor protein TGR17 of human origin and DNA
PT encoding it for diagnosis and treatment of cancer and circulatory and
PT other diseases associated with its expression -
XX
PS Claim 3; Fig 12; 145pp; Japanese.
XX
CC The present invention describes a human guanine nucleotide binding
CC protein (G protein) coupled receptor protein designated TGR17, which has
CC 5 variant forms (TGR17-1 to TGR17-5) of the receptor protein having
CC additional residues at the N-terminal. The TGR17 proteins have nootropic,
CC antiinflammatory, vasotropic, immunomodulator and cytostatic activities.
CC The TGR17 polynucleotides and protein can be used in gene therapy and
CC protein therapy. G-protein coupled receptor proteins are cell membrane
CC proteins mediating the cellular response to a large variety of signalling
CC molecules. The TGR17 polynucleotides and proteins can be used in the
CC diagnosis, treatment and prevention of diseases including neurological,
CC inflammatory, circulatory, degenerative, immune system and digestive
CC diseases and cancer. The present sequence represents human TGR17-6 from
CC the present invention.
XX
SQ Sequence 730 AA;

Query Match 100.0%; Score 549; DB 23; Length 730;
Best Local Similarity 100.0%; Pred. No. 4.1e-65;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 572 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 631

Qy 61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
|||||||||||||||||||||||||||||||||||||||||||||
Db 632 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 678

RESULT 11
AAU97158
ID AAU97158 standard; Protein; 737 AA.
XX
AC AAU97158;
XX
DT 27-AUG-2002 (first entry)
XX
DE Human G-protein coupled receptor HGPRBMY5.
XX
KW Human; G-protein coupled receptor; GPCR; HGPRBMY5; colon; brain;
KW ovary; thymus; lung; immune system; cancer; immune disorder;
KW neurological disorder; infection; human immunodeficiency virus; HIV;
KW antiallergic; antiasthmatic; dermatological; antiarteriosclerotic;
KW antiulcer; antidiabetic; nephrotropic; osteopathic; antiarthritic;
KW antiinflammatory; antirheumatic; antithyroid; cytostatic; vulnerary;
KW virucide; antibacterial; antifungal; antiparasitic; protozoacide;
KW antihelminthic; nootropic; neuroprotective; antidepressant;
KW anticonvulsant; antiparkinsonian; neuroleptic; anti-HIV; receptor.
XX
OS Homo sapiens.
XX
PN WO200226824-A2.
XX

PD 04-APR-2002.
XX
PF 26-SEP-2001; 2001WO-US30365.
XX
PR 27-SEP-2000; 2000US-235713P.
PR 16-JAN-2001; 2001US-261781P.
PR 19-JUL-2001; 2001US-306605P.
PR 03-AUG-2001; 2001US-310436P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Feder JN, Mintier G, Ramanathan CS, Hawken DR;
XX
DR WPI; 2002-435196/46.
DR N-PSDB; ABK51944.
XX
PT Novel G protein-coupled receptor, HGPRBMY5 polypeptide, useful for
PT treatment of AIDS, allergies, asthma, atherosclerosis, ulcerative
PT colitis, atopic dermatitis, diabetes mellitus, glomerulonephritis,
PT osteoarthritis -
XX
PS Claim 11; Fig 2; 148pp; English.
XX
CC The present invention relates to the isolation of a novel human
CC G-protein coupled receptor (GPCR) (HGPRBMY5), and the polynucleotide
CC sequence encoding it. The HGPRBMY5 polypeptide and polynucleotide
CC are useful for preventing, treating or ameliorating a disease,
CC disorder or condition related to the colon, brain, ovaries, thymus,
CC lungs or imm immune system. They are particularly useful for the
CC treatment or prevention of cancers, immune disorders, neurological
CC disorders, and diseases related to the brain, ovaries, thymus or
CC lungs. The polynucleotide sequence is useful for diagnosing or
CC determining susceptibility to infections such as bacterial, fungal,
CC protozoan and viral infections, particularly infections caused by
CC human immunodeficiency virus (HIV or HIV-2). The present sequence
CC represents human GPCR HGPRBMY5.
XX
SQ Sequence 737 AA;

Query Match 100.0%; Score 549; DB 23; Length 737;
Best Local Similarity 100.0%; Pred. No. 4.1e-65;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 579 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 638

Qy 61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
|||||||||||||||||||||||||||||||||||||||||||||
Db 639 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 685

RESULT 12
ABB06254
ID ABB06254 standard; Protein; 754 AA.
XX
AC ABB06254;
XX
DT 23-MAY-2002 (first entry)
XX
DE Human G protein-coupled receptor TGR17-5 protein SEQ ID NO:13.
XX
KW Human; G protein-coupled receptor; TGR17-5; nootropic; antiinflammatory;
KW vasotropic; immunomodulator; cytostatic; gene therapy; protein therapy;
KW neurological; inflammatory; circulatory; degenerative; immune system;
KW digestive disease; cancer.
XX
OS Homo sapiens.
XX
PN WO200204640-A1.
XX
PD 17-JAN-2002.
XX
PF 06-JUL-2001; 2001WO-JP05878.
XX
PR 07-JUL-2000; 2000JP-0211989.
PR 18-DEC-2000; 2000JP-0383794.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Moriya T, Ito T, Shintani Y, Miyajima N;
XX
DR WPI; 2002-179706/23.
DR N-PSDB; ABL40196.
XX
PT G-protein coupled receptor protein TGR17 of human origin and DNA
PT encoding it for diagnosis and treatment of cancer and circulatory and
PT other diseases associated with its expression -
XX
PS Claim 3; Fig 11; 145pp; Japanese.
XX
CC The present invention describes a human guanine nucleotide binding
CC protein (G protein) coupled receptor protein designated TGR17, which has
CC 5 variant forms (TGR17-1 to TGR17-5) of the receptor protein having
CC additional residues at the N-terminal. The TGR17 proteins have nootropic,
CC antiinflammatory, vasotropic, immunomodulator and cytostatic activities.
CC The TGR17 polynucleotides and protein can be used in gene therapy and
CC protein therapy. G-protein coupled receptor proteins are cell membrane
CC proteins mediating the cellular response to a large variety of signalling
CC molecules. The TGR17 polynucleotides and proteins can be used in the
CC diagnosis, treatment and prevention of diseases including neurological,
CC inflammatory, circulatory, degenerative, immune system and digestive
CC diseases and cancer. The present sequence represents human TGR17-5 from
CC the present invention.
XX
SQ Sequence 754 AA;

Query Match 100.0%; Score 549; DB 23; Length 754;
Best Local Similarity 100.0%; Pred. No. 4.3e-65;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 596 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 655

Qy 61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
|||||||||||||||||||||||||||||||||||||||||||||
Db 656 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 702

RESULT 13
AAU87545
ID AAU87545 standard; Protein; 176 AA.
XX
AC AAU87545;
XX
DT 05-JUN-2002 (first entry)
XX
DE Novel central nervous system protein #455.
XX
KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;
KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
KW adenocarcinoma; reproductive system disorder; testicular feminisation;
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
KW respiratory disorder; renal disorder; kidney failure; blood disorder;
KW myocardial infarction; wound healing; cell proliferation; skin aging;
KW food additive; food preservative; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200155318-A2.
XX
PD 02-AUG-2001.

XX
PF 17-JAN-2001; 2001WO-US01332.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-581633/65.
DR N-PSDB; ABK43875.

```
XX
PT   New isolated nucleic acid encoding a protein for diagnosing,
PT   preventing, treating or ameliorating medical conditions and used as
PT   food additives or preservatives -
XX
PS   Claim 9; SEQ ID No 1063; 837pp; English.
XX
CC   The invention describes an isolated nucleic acid molecule (I) encoding a
CC   novel central nervous system protein. (I) and polypeptides (III) encoded
CC   by (I), are used to treat a medical conditions and in diagnosis of a
CC   pathological condition. Disorders which are diagnosed or treated include
CC   autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC   disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC   e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC   angiogenesis, nervous system disorders e.g. Alzheimer's disease and
CC   amylotrophic lateral sclerosis, infections caused by bacteria, viruses
CC   e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
CC   e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
CC   adenocarcinomas and irritable bowel syndrome, reproductive system
CC   disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
CC   and pituitary dwarfism, cancers and disorders at the cellular level e.g.
CC   leukaemia, disorders involving neovascularisation e.g. malignancies,
CC   respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
CC   acute kidney failure and blood related disorders e.g. myocardial
CC   infarction. The polypeptides can also be used to aid wound healing and
CC   epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC   maintain organs before transplantation, for supporting cell culture of
CC   primary tissues, to regenerate tissues and in chemotaxis. The
CC   polypeptides can also be used as a food additive or preservative to
CC   increase or decrease storage capabilities, fat content, lipid, protein,

  Query Match             68.5%;  Score 376;  DB 22;  Length 176;
  Best Local Similarity   64.5%;  Pred. No. 1.6e-42;
  Matches   69;  Conservative   20;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
           |:|| ||:|||||| :|| |: ::|:  ||:||   :|: :| |||||||:||:||||:|
Db      11 GINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVFTDALCWIPIF 70

Qy      61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
           ||| ||| :|||| |:|||:||| ||:||||||||||||  ||: :
Db      71 VVKFLSLLQVEIPGTITSWVVIFILPINSALNPILYTLTTRPFKEMI 117


RESULT 14
AAM99972
ID   AAM99972 standard; Protein; 176 AA.
XX
AC   AAM99972;
XX
DT   04-JAN-2002  (first entry)
XX
DE   Human expressed polypeptide SEQ ID NO 96.
XX
KW   Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW   immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW   antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW   antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW   antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW   antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; nephrotropic; gene therapy; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200155387-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01310.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
```

PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465573/50.
DR N-PSDB; AAI99584.
XX
PT Isolated digestive system associated polypeptide for treating,
PT preventing and/ or prognosing disorders related to the digestive system
PT including digestive system cancers and also for testing and detection
PT e.g. diagnosis -
XX
PS Claim 11; SEQ ID NO 96; 509pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (AAI99548-AAI99604) and proteins
CC (AAM99936-AAM99984) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 176 AA;

Query Match 68.5%; Score 376; DB 22; Length 176;
Best Local Similarity 64.5%; Pred. No. 1.6e-42;
Matches 69; Conservative 20; Mismatches 18; Indels 0; Gaps 0;

```
Qy       1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
           |:|| ||:|||||| :|| |: ::|:  ||:||   :|: :| |||||||:||:||||:|
Db      11 GINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVFTDALCWIPIF 70

Qy      61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
           ||| ||| :|||| |:|||:||| ||:||||||||||||||  ||: :
Db      71 VVKFLSLLQVEIPGTITSWVVIFILPINSALNPILYTLTTRPFKEMI 117
```

RESULT 15
ABB04062
ID ABB04062 standard; Protein; 176 AA.
XX
AC ABB04062;
XX
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polypeptide SEQ ID NO 2009.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system.
XX
OS Homo sapiens.
XX
PN WO200155367-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01338.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-451937/48.
DR N-PSDB; AAL35644.
XX
PT Isolated polypeptide for treating, preventing and/ or prognosing
PT disorders related to the musculoskeletal system including
PT musculoskeletal cancers and also for testing and detection e.g.
PT diagnosis -
XX
PS Claim 11; SEQ ID NO 2009; 781pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (AAL34669-AAL37666) and proteins
CC (ABB03087-ABB04109) associated with the musculoskeletal system useful
CC for preventing, treating or ameliorating medical conditions e.g. by
CC protein or gene therapy. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids, proteins,
CC antibodies and (ant)agonists are useful in the diagnosis, treatment
CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and

CC other cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC (c) cardiovascular disorders such as myocardial ischaemias; (d) wound
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC and (f) infectious diseases such as viral, bacterial, fungal and
CC parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 176 AA;

Query Match 68.5%; Score 376; DB 22; Length 176;
Best Local Similarity 64.5%; Pred. No. 1.6e-42;
Matches 69; Conservative 20; Mismatches 18; Indels 0; Gaps 0;

```
Qy       1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
           |:|| ||:|||||| :|| |: ::|:  ||:||   :|: :| |||||||:||:||||:|
Db      11 GINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVFTDALCWIPIF 70

Qy      61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
           ||| ||| :|||| |:|||:||| ||:||||||||||||  ||: :
Db      71 VVKFLSLLQVEIPGTITSWVVIFILPINSALNPILYTLTTRPFKEMI 117
```

Search completed: May 22, 2003, 12:45:00
Job time : 76 secs

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 22, 2003, 12:43:45 ; Search time 29 Seconds
(without alignments)
108.560 Million cell updates/sec

Title: US-09-930-312-2
Perfect score: 549
Sequence: 1 GVNLLAFLIIVFSYITMFCS.....NSALNPILYTLTTNFFKDKL 107

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	142.5	26.0	792	4	US-07-741-453A-56	Sequence 56, Appl
2	139.5	25.4	696	4	US-07-757-342D-4	Sequence 4, Appli
3	138.5	25.2	320	1	US-08-101-435-4	Sequence 4, Appli
4	138.5	25.2	611	4	US-07-757-342D-8	Sequence 8, Appli
5	138.5	25.2	636	4	US-07-757-342D-7	Sequence 7, Appli
6	138.5	25.2	674	4	US-07-757-342D-10	Sequence 10, Appl
7	138.5	25.2	699	4	US-07-757-342D-2	Sequence 2, Appli
8	136.5	24.9	700	4	US-07-757-342D-3	Sequence 3, Appli
9	129.5	23.6	795	4	US-07-741-453A-55	Sequence 55, Appl
10	127.5	23.2	332	1	US-08-118-270-53	Sequence 53, Appl
11	127.5	23.2	332	5	PCT-US93-08528-53	Sequence 53, Appl
12	125	22.8	420	4	US-08-795-876-33	Sequence 33, Appl
13	125	22.8	423	4	US-08-795-876-38	Sequence 38, Appl
14	125	22.8	436	4	US-08-795-876-2	Sequence 2, Appli
15	125	22.8	695	1	US-08-487-886-2	Sequence 2, Appli
16	125	22.8	695	3	US-08-482-855-2	Sequence 2, Appli
17	125	22.8	695	4	US-08-474-986-2	GENERAL INFORMA
18	123	22.4	194	1	US-07-918-314-4	Sequence 4, Appli
19	123	22.4	326	1	US-07-918-314-6	Sequence 6, Appli
20	123	22.4	326	1	US-08-349-696-19	Sequence 19, Appl
21	123	22.4	326	1	US-08-233-009-19	Sequence 19, Appl
22	123	22.4	326	2	US-08-560-231-19	Sequence 19, Appl
23	123	22.4	326	4	US-09-080-704A-19	Sequence 19, Appl
24	122.5	22.3	764	4	US-07-741-453A-60	Sequence 60, Appl
25	122	22.2	692	4	US-07-757-342D-6	Sequence 6, Appli
26	120.5	21.9	764	4	US-07-741-453A-54	Sequence 54, Appl
27	120	21.9	326	1	US-08-293-563-7	Sequence 7, Appli
28	119.5	21.8	318	1	US-08-349-696-25	Sequence 25, Appl
29	119.5	21.8	318	1	US-08-233-009-25	Sequence 25, Appl
30	119.5	21.8	318	2	US-08-560-231-25	Sequence 25, Appl
31	119.5	21.8	318	4	US-09-080-704A-25	Sequence 25, Appl
32	119.5	21.8	336	1	US-08-118-270-54	Sequence 54, Appl
33	119.5	21.8	336	5	PCT-US93-08528-54	Sequence 54, Appl
34	119.5	21.8	764	4	US-07-741-453A-59	Sequence 59, Appl
35	119.5	21.8	764	4	US-07-741-453A-61	Sequence 61, Appl
36	118.5	21.6	764	4	US-07-757-342D-5	Sequence 5, Appli
37	116	21.1	332	1	US-08-293-563-5	Sequence 5, Appli
38	115	20.9	332	1	US-08-349-696-23	Sequence 23, Appl
39	115	20.9	332	2	US-08-560-231-23	Sequence 23, Appl
40	115	20.9	332	4	US-09-080-704A-23	Sequence 23, Appl
41	113.5	20.7	334	1	US-08-118-270-73	Sequence 73, Appl
42	113.5	20.7	334	5	PCT-US93-08528-73	Sequence 73, Appl
43	110	20.0	332	1	US-08-233-009-23	Sequence 23, Appl
44	109	19.9	644	2	US-08-866-757-2	Sequence 2, Appli
45	109	19.9	644	4	US-09-153-593-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-07-741-453A-56
; Sequence 56, Application US/07741453A
; Patent No. 6228597
; GENERAL INFORMATION:
; APPLICANT: PARMENTIER, MARC
; APPLICANT: LIBERT, FREDERIC
; APPLICANT: DUMONT, JACQUES
; APPLICANT: VASSART, GILBERT
; TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR
; TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS
; TITLE OF INVENTION: AND POLYPEPTIDES, AND APPLICATIONS OF THESE POLYPEPTIDES
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1615 L STREET, N.W.
; CITY: WASHINGTON, D.C.
; COUNTRY: U.S.A.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/741,453A
; FILING DATE: 19911015
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16773
; REFERENCE/DOCKET NUMBER: 91913/1107/US/ST
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 792 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-741-453A-56

Query Match 26.0%; Score 142.5; DB 4; Length 792;
Best Local Similarity 27.9%; Pred. No. 4.5e-08;
Matches 31; Conservative 24; Mismatches 43; Indels 13; Gaps 2;

Qy 1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60

```
          |:|::||:||   || :: ::|   |  |       ::  :| :   ::|:|  |  |:
Db    611 GLNVVAFIIICACYIKIYFAVQNPELMAT------NKDTKIAKKMAVLIFTDFTCMAPIG 664

Qy     61 VVKILSLFRVE-------IPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
          :| | :          |  | :  ::: | ||||  || || : |  |:
Db    665 HCGPISFFAISAALKVPLITVTNSKVLLVLFYPVNSCANPFLYAIFTKAFR 715


RESULT 2
US-07-757-342D-4
; Sequence 4, Application US/07757342D
; Patent No. 6218509
;   GENERAL INFORMATION:
;        APPLICANT: IGARASHI, Masao
;                   MINEGISHI, Takashi
;                   NAKAMURA, Kazuto
;        TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
;        NUMBER OF SEQUENCES: 10
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
;                        CUSHMAN
;             STREET: 130 Water Street
;             CITY: Boston
;             STATE: Massachusetts
;             COUNTRY: US
;             ZIP: 02109
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/07/757,342D
;             FILING DATE: 10-Sep-1991
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BUCKLEY, Linda M.
;             REGISTRATION NUMBER: 31003
;             REFERENCE/DOCKET NUMBER: 41226
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617)523-3400
;             TELEFAX: (617)523-6440
;             TELEX: 200291 STRE UR
;   INFORMATION FOR SEQ ID NO: 4:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 696 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-07-757-342D-4

  Query Match             25.4%;  Score 139.5;  DB 4;  Length 696;
  Best Local Similarity   28.8%;  Pred. No. 8.5e-08;
  Matches   30;  Conservative   24;  Mismatches   43;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|::||:||   || :: ::|   |  |       ::  :| :   ::|:|  |  |:
Db     534 LNVVAFIIICACYIKIYFAVQNPELMAT------NKDTKIAKKMAVLIFTDFTCMAPISF 587

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
           | :  :|  |  | :  ::: | ||||  || || : |  |:
Db     588 FAISAALKVPLITVTNSKVLLVLFYPVNSCANPFLYAIFTKAFR 631


RESULT 3
US-08-101-435-4
; Sequence 4, Application US/08101435
; Patent No. 5441883
;  GENERAL INFORMATION:
;    APPLICANT:  Civelli, Olivier
;    APPLICANT:  Zhou, Qun-yong
;    TITLE OF INVENTION:  A No. 5441883el Adenosine Receptor and Uses
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Allegretti & Witcoff, Ltd.
;      STREET:  10 South Wacker Drive, Suite 3000
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/101,435
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/847,563
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McDonnell, John J
;      REGISTRATION NUMBER:  26,949
;      REFERENCE/DOCKET NUMBER:  91,708
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-715-1000
;      TELEFAX:  312-715-1234
;      TELEX:  910-221-5317
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  320 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-101-435-4

  Query Match             25.2%;  Score 138.5;  DB 1;  Length 320;
  Best Local Similarity   29.9%;  Pred. No. 4.7e-08;
  Matches   29;  Conservative   22;  Mismatches   41;  Indels    5;  Gaps    2;

Qy       4 LLAFLIIVFSYITMFCSIQKTALQT----TEVRNCFGREVAVANRFFFIVFSDAICWIPV 59
           |:  :::   |: :|  |:     |      | |  :|||   |   | ::|  |:||:|:
Db     189 LIPLVVMCIIYLDIFYIIRNKLSQNLTGFRETRAFYGREFKTAKSLFLVLFLFALCWLPL 248

Qy      60 FVVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILY 96
            ::  :| | |:||: :   : |     || :|||:|
Db     249 SIINFVSYFNVKIPE-IAMCLGILLSHANSMMNPIVY 284


RESULT 4
US-07-757-342D-8
; Sequence 8, Application US/07757342D
; Patent No. 6218509
;   GENERAL INFORMATION:
;        APPLICANT: IGARASHI, Masao
;                   MINEGISHI, Takashi
;                   NAKAMURA, Kazuto
;        TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
;        NUMBER OF SEQUENCES: 10
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
;                        CUSHMAN
;             STREET: 130 Water Street
;             CITY: Boston
;             STATE: Massachusetts
;             COUNTRY: US
;             ZIP: 02109
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
```

```
;             APPLICATION NUMBER: US/07/757,342D
;             FILING DATE: 10-Sep-1991
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BUCKLEY, Linda M.
;             REGISTRATION NUMBER: 31003
;             REFERENCE/DOCKET NUMBER: 41226
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617)523-3400
;             TELEFAX: (617)523-6440
;             TELEX: 200291 STRE UR
;   INFORMATION FOR SEQ ID NO: 8:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 611 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-07-757-342D-8

  Query Match             25.2%;  Score 138.5;  DB 4;  Length 611;
  Best Local Similarity   27.9%;  Pred. No. 9.6e-08;
  Matches   29;  Conservative   25;  Mismatches   43;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|::|| ||   || :: :::   |  |       ::  :| :   ::|:|  |  |:
Db     446 LNVVAFFIICACYIKIYFAVRNPELMAT------NKDTKIAKKMAILIFTDFTCMAPISF 499

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
             | : |:|  |  | :  ::: | |:||  || || : |  |:
Db     500 FAISAAFKVPLITVTNSKVLLVLFYPINSCANPFLYAIFTKTFQ 543


RESULT 5
US-07-757-342D-7
; Sequence 7, Application US/07757342D
; Patent No. 6218509
;   GENERAL INFORMATION:
;        APPLICANT: IGARASHI, Masao
;                   MINEGISHI, Takashi
;                   NAKAMURA, Kazuto
;        TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
;        NUMBER OF SEQUENCES: 10
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
;                        CUSHMAN
;             STREET: 130 Water Street
;             CITY: Boston
;             STATE: Massachusetts
;             COUNTRY: US
;             ZIP: 02109
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/07/757,342D
;             FILING DATE: 10-Sep-1991
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BUCKLEY, Linda M.
;             REGISTRATION NUMBER: 31003
;             REFERENCE/DOCKET NUMBER: 41226
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617)523-3400
;             TELEFAX: (617)523-6440
;             TELEX: 200291 STRE UR
;   INFORMATION FOR SEQ ID NO: 7:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 636 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-07-757-342D-7

  Query Match             25.2%;  Score 138.5;  DB 4;  Length 636;
  Best Local Similarity   27.9%;  Pred. No. 1e-07;
  Matches   29;  Conservative   25;  Mismatches   43;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|::|| ||   || :: :::   |  |       ::  :| :   ::|:|  |  |:
Db     471 LNVVAFFIICACYIKIYFAVRNPELMAT------NKDTKIAKKMAILIFTDFTCMAPISF 524

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
             | : |:|  |  | :  ::: | |:||  || || : |  |:
Db     525 FAISAAFKVPLITVTNSKVLLVLFYPINSCANPFLYAIFTKTFQ 568


RESULT 6
US-07-757-342D-10
; Sequence 10, Application US/07757342D
; Patent No. 6218509
;   GENERAL INFORMATION:
;        APPLICANT: IGARASHI, Masao
;                   MINEGISHI, Takashi
;                   NAKAMURA, Kazuto
;        TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
;        NUMBER OF SEQUENCES: 10
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
;                        CUSHMAN
;             STREET: 130 Water Street
;             CITY: Boston
;             STATE: Massachusetts
;             COUNTRY: US
;             ZIP: 02109
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/07/757,342D
;             FILING DATE: 10-Sep-1991
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BUCKLEY, Linda M.
;             REGISTRATION NUMBER: 31003
;             REFERENCE/DOCKET NUMBER: 41226
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617)523-3400
;             TELEFAX: (617)523-6440
;             TELEX: 200291 STRE UR
;   INFORMATION FOR SEQ ID NO: 10:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 674 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-07-757-342D-10

  Query Match             25.2%;  Score 138.5;  DB 4;  Length 674;
  Best Local Similarity   27.9%;  Pred. No. 1.1e-07;
  Matches   29;  Conservative   25;  Mismatches   43;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|::|| ||   || :: :::   |  |       ::  :| :   ::|:|  |  |:
Db     509 LNVVAFFIICACYIKIYFAVRNPELMAT------NKDTKIAKKMAILIFTDFTCMAPISF 562

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
             | : |:|  |  | :  ::: | |:||  || || : |  |:
Db     563 FAISAAFKVPLITVTNSKVLLVLFYPINSCANPFLYAIFTKTFQ 606
```

```
RESULT 7
US-07-757-342D-2
; Sequence 2, Application US/07757342D
; Patent No. 6218509
;  GENERAL INFORMATION:
;       APPLICANT: IGARASHI, Masao
;                  MINEGISHI, Takashi
;                  NAKAMURA, Kazuto
;       TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
;       NUMBER OF SEQUENCES: 10
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
;                       CUSHMAN
;            STREET: 130 Water Street
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: US
;            ZIP: 02109
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/07/757,342D
;            FILING DATE: 10-Sep-1991
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: BUCKLEY, Linda M.
;            REGISTRATION NUMBER: 31003
;            REFERENCE/DOCKET NUMBER: 41226
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)523-3400
;            TELEFAX: (617)523-6440
;            TELEX: 200291 STRE UR
;  INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 699 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-07-757-342D-2

  Query Match             25.2%;  Score 138.5;  DB 4;  Length 699;
  Best Local Similarity   27.9%;  Pred. No. 1.1e-07;
  Matches   29;  Conservative   25;  Mismatches   43;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|::||  ||   || :: :::   | |       ::  :| :   ::|:|  |  |:
Db     534 LNVVAFFIICACYIKIYFAVRNPELMAT------NKDTKIAKKMAILIFTDFTCMAPISF 587

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
             | : |:|  |  | :  ::: | |:||  || || : |  |:
Db     588 FAISAAFKVPLITVTNSKVLLVLFYPINSCANPFLYAIFTKTFQ 631


RESULT 8
US-07-757-342D-3
; Sequence 3, Application US/07757342D
; Patent No. 6218509
;  GENERAL INFORMATION:
;       APPLICANT: IGARASHI, Masao
;                  MINEGISHI, Takashi
;                  NAKAMURA, Kazuto
;       TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
;       NUMBER OF SEQUENCES: 10
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
;                       CUSHMAN
;            STREET: 130 Water Street
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: US
;            ZIP: 02109
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/07/757,342D
;            FILING DATE: 10-Sep-1991
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: BUCKLEY, Linda M.
;            REGISTRATION NUMBER: 31003
;            REFERENCE/DOCKET NUMBER: 41226
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)523-3400
;            TELEFAX: (617)523-6440
;            TELEX: 200291 STRE UR
;  INFORMATION FOR SEQ ID NO: 3:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 700 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-07-757-342D-3

  Query Match             24.9%;  Score 136.5;  DB 4;  Length 700;
  Best Local Similarity   27.9%;  Pred. No. 1.9e-07;
  Matches   29;  Conservative   25;  Mismatches   43;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|::||::|   || :: ::|   |          ::  :| :   ::|:|  |  |:
Db     538 LNVVAFVVICACYIRIYFAVQNPELTAP------NKDTKIAKKMAILIFTDFTCMAPISF 591

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
             | : |:|  |  | :  ::: | ||||  || || : |  |:
Db     592 FAISAAFKVPLITVTNSKILLVLFYPVNSCANPFLYAIFTKAFQ 635


RESULT 9
US-07-741-453A-55
; Sequence 55, Application US/07741453A
; Patent No. 6228597
;  GENERAL INFORMATION:
;    APPLICANT:  PARMENTIER, MARC
;    APPLICANT:  LIBERT, FREDERIC
;    APPLICANT:  DUMONT, JACQUES
;    APPLICANT:  VASSART, GILBERT
;    TITLE OF INVENTION:  POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR
;    TITLE OF INVENTION:  ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS
;    TITLE OF INVENTION:  AND POLYPEPTIDES, AND APPLICATIONS OF THESE POLYPEPTIDES
;    NUMBER OF SEQUENCES:  62
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CUSHMAN DARBY & CUSHMAN
;      STREET:  1615 L STREET, N.W.
;      CITY:  WASHINGTON, D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/741,453A
;      FILING DATE:  19911015
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KOKULIS, PAUL N.
;      REGISTRATION NUMBER:.  16773
```

```
;      REFERENCE/DOCKET NUMBER:  91913/1107/US/ST
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 861-3000
;      TELEFAX:  (202) 822-0944
;      TELEX:  6714627 CUSH
;  INFORMATION FOR SEQ ID NO:  55:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  795 amino acids
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-741-453A-55

  Query Match             23.6%;  Score 129.5;  DB 4;  Length 795;
  Best Local Similarity   27.0%;  Pred. No. 1.3e-06;
  Matches   30;  Conservative   25;  Mismatches   43;  Indels   13;  Gaps    3;

Qy       1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWI--- 57
           |:|::||::|   || :: ::|   |          ::  :| :   ::|:|  |
Db     614 GLNVVAFVVICACYIRIYFAVQNPELTAP------NKDTKIAKKMAILIFTDFTCMARAT 667

Qy      58 ---PVFVVKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
             |:    | : |:|  |  | :  ::: | ||||  || || : |  |:
Db     668 HCGPISFFAISAAFKVPLITVTNSKILLVLFYPVNSCANPFLYAIFTKAFQ 718


RESULT 10
US-08-118-270-53
; Sequence 53, Application US/08118270
; Patent No. 5508384
;  GENERAL INFORMATION:
;    APPLICANT:  Murphy, Randall B.
;    APPLICANT:  Schuster, David I.
;    TITLE OF INVENTION:  POLYPEPTIDES OF G-COUPLED PROTEIN
;    TITLE OF INVENTION:  RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
;    NUMBER OF SEQUENCES:  348
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/118,270
;      FILING DATE:  09-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/943,236
;      FILING DATE:  10-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Townsend, Kevin G.
;      REGISTRATION NUMBER:  34,033
;      REFERENCE/DOCKET NUMBER:  MURPHY=2A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  53:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  332 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-118-270-53

  Query Match             23.2%;  Score 127.5;  DB 1;  Length 332;
  Best Local Similarity   28.2%;  Pred. No. 8.3e-07;
  Matches   29;  Conservative   27;  Mismatches   38;  Indels    9;  Gaps    3;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|::|||||   || :: :::   :  |        ::|:|     ::|:|  | :|:
Db     173 LNVVAFLIICACYIKIYFAVRNPEIMATNK----DTKIALA----ILIFTDFTC-MPISF 223

Qy      62 VKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
             | : |:| :  | :  ::: | |:||  || || : |  |:
Db     224 FAISAAFKVPLIVTNSKVLLVLFYPINSCANPFLYAIFTKTFQ 266


RESULT 11
PCT-US93-08528-53
; Sequence 53, Application PC/TUS9308528
;  GENERAL INFORMATION:
;    APPLICANT:  New York University
;    TITLE OF INVENTION:  POLYPEPTIDES OF G-COUPLED PROTEIN
;    TITLE OF INVENTION:  RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
;    NUMBER OF SEQUENCES:  348
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/08528
;      FILING DATE:  09-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/943,236
;      FILING DATE:  10-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Townsend, Kevin G.
;      REGISTRATION NUMBER:  34,033
;      REFERENCE/DOCKET NUMBER:  MURPHY=2 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  53:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  332 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
PCT-US93-08528-53

  Query Match             23.2%;  Score 127.5;  DB 5;  Length 332;
  Best Local Similarity   28.2%;  Pred. No. 8.3e-07;
  Matches   29;  Conservative   27;  Mismatches   38;  Indels    9;  Gaps    3;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|::|||||   || :: :::   :  |        ::|:|     ::|:|  | :|:
Db     173 LNVVAFLIICACYIKIYFAVRNPEIMATNK----DTKIALA----ILIFTDFTC-MPISF 223

Qy      62 VKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
             | : |:| :  | :  ::: | |:||  || || : |  |:
Db     224 FAISAAFKVPLIVTNSKVLLVLFYPINSCANPFLYAIFTKTFQ 266


RESULT 12
US-08-795-876-33
; Sequence 33, Application US/08795876
; Patent No. 6403305
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Gershengorn, Marvin C.
;    APPLICANT:  Geras-Raaka, Elizabeth
;    APPLICANT:  Nussenzveig, Daniel R.
;    TITLE OF INVENTION:  STRATEGY TO CLONE DRUGS FOR G PROTEIN
;    TITLE OF INVENTION:  COUPLED RECEPTORS
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON, HARGRAVE, DEVANS & DOYLE LLP
;      STREET:  Clinton Square, P.O. Box 1051
;      CITY:  Rochester
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  14603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/795,876
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BRAMAN, SUSAN J.
;      REGISTRATION NUMBER:  34,103
;      REFERENCE/DOCKET NUMBER:  19603/1280
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  716-263-1636
;      TELEFAX:  716-263-1600
;  INFORMATION FOR SEQ ID NO:  33:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  420 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-795-876-33

  Query Match             22.8%;  Score 125;  DB 4;  Length 420;
  Best Local Similarity   26.4%;  Pred. No. 2.1e-06;
  Matches   28;  Conservative   28;  Mismatches   42;  Indels    8;  Gaps    3;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|:|||::|   || :: :::   : ::        :  :| |   ::|:| :|  |:
Db     262 LNVLAFVVICGCYIHIYLTVRNPNIVSS------SSDTRIAKRMAMLIFTDFLCMAPISF 315

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTL-TTNFFKD 105
             | :  :|  |  :    ::: | |:||  || || : | || :|
Db     316 FAISASLKVPLITVSKAKILLVLFHPINSCANPFLYAIFTKNFRRD 361


RESULT 13
US-08-795-876-38
; Sequence 38, Application US/08795876
; Patent No. 6403305
;  GENERAL INFORMATION:
;    APPLICANT:  Gershengorn, Marvin C.
;    APPLICANT:  Geras-Raaka, Elizabeth
;    APPLICANT:  Nussenzveig, Daniel R.
;    TITLE OF INVENTION:  STRATEGY TO CLONE DRUGS FOR G PROTEIN
;    TITLE OF INVENTION:  COUPLED RECEPTORS
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON, HARGRAVE, DEVANS & DOYLE LLP
;      STREET:  Clinton Square, P.O. Box 1051
;      CITY:  Rochester
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  14603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/795,876
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BRAMAN, SUSAN J.
;      REGISTRATION NUMBER:  34,103
;      REFERENCE/DOCKET NUMBER:  19603/1280
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  716-263-1636
;      TELEFAX:  716-263-1600
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  423 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-795-876-38

  Query Match             22.8%;  Score 125;  DB 4;  Length 423;
  Best Local Similarity   26.4%;  Pred. No. 2.1e-06;
  Matches   28;  Conservative   28;  Mismatches   42;  Indels    8;  Gaps    3;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|:|||::|   || :: :::   : ::        :  :| |   ::|:| :|  |:
Db     265 LNVLAFVVICGCYIHIYLTVRNPNIVSS------SSDTRIAKRMAMLIFTDFLCMAPISF 318

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTL-TTNFFKD 105
             | :  :|  |  :    ::: | |:||  || || : | || :|
Db     319 FAISASLKVPLITVSKAKILLVLFHPINSCANPFLYAIFTKNFRRD 364


RESULT 14
US-08-795-876-2
; Sequence 2, Application US/08795876
; Patent No. 6403305
;  GENERAL INFORMATION:
;    APPLICANT:  Gershengorn, Marvin C.
;    APPLICANT:  Geras-Raaka, Elizabeth
;    APPLICANT:  Nussenzveig, Daniel R.
;    TITLE OF INVENTION:  STRATEGY TO CLONE DRUGS FOR G PROTEIN
;    TITLE OF INVENTION:  COUPLED RECEPTORS
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON, HARGRAVE, DEVANS & DOYLE LLP
;      STREET:  Clinton Square, P.O. Box 1051
;      CITY:  Rochester
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  14603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/795,876
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BRAMAN, SUSAN J.
;      REGISTRATION NUMBER:  34,103
;      REFERENCE/DOCKET NUMBER:  19603/1280
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  716-263-1636
;      TELEFAX:  716-263-1600
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  436 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
```

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-795-876-2

  Query Match             22.8%;  Score 125;  DB 4;  Length 436;
  Best Local Similarity   26.4%;  Pred. No. 2.1e-06;
  Matches   28;  Conservative   28;  Mismatches   42;  Indels    8;  Gaps    3;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|:|||::|   || :: :::   : ::       :  :| |   ::|:| :|  |:
Db     278 LNVLAFVVICGCYIHIYLTVRNPNIVSS------SSDTRIAKRMAMLIFTDFLCMAPISF 331

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTL-TTNFFKD 105
             |  :  :|  |  :    ::: | |:||  || || : | || :|
Db     332 FAISASLKVPLITVSKAKILLVLFHPINSCANPFLYAIFTKNFRRD 377


RESULT 15
US-08-487-886-2
; Sequence 2, Application US/08487886
; Patent No. 5744448
;  GENERAL INFORMATION:
;    APPLICANT:  Kelton, Christie Ann
;    APPLICANT:  Schweickhardt, Rene Lynn
;    APPLICANT:  Cheng, Shirley Vui Yen
;    APPLICANT:  Nugent, No. 5744448een Patrice
;    TITLE OF INVENTION:  Human Follicle Stimulating
;    TITLE OF INVENTION:  Hormone Receptor
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Stephan P. Williams,
;      ADDRESSEE:  Ares-Serono, Inc.
;      STREET:  Exchange Place, 37th floor
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" diskette, 1.44 MB, high density
;      COMPUTER:  IBM PS/2, model 55 SX
;      OPERATING SYSTEM:  MS-DOS version 4.0
;      SOFTWARE:  VAX/VMS Mass11 via Kermit to IBM MS-DOS
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,886
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/670,085
;      FILING DATE:  15-MAR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams, Stephan P.
;      REGISTRATION NUMBER:  28546
;      REFERENCE/DOCKET NUMBER:  US/252
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 723-1300
;      TELEFAX:  (617) 723-8923
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  695
;      TYPE:  Amino acid
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  signal sequence
;      LOCATION:  -17 to -1
;      IDENTIFICATION METHOD:  hydrophobic
;    FEATURE:
;      NAME/KEY:  putative amino-terminal extracellular domain
;      LOCATION:  1 to 349
;      IDENTIFICATION METHOD:  similarity with other
;      IDENTIFICATION METHOD:  dimeric glycoprotein receptor extracellular
;      IDENTIFICATION METHOD:  domains, hydrophilic
;    FEATURE:
;      NAME/KEY:  transmembrane domain
;      LOCATION:  350 to 613
;      IDENTIFICATION METHOD:  similarity to other G
;      IDENTIFICATION METHOD:  protein-coupled receptor transmembrane domains
;    FEATURE:
;      NAME/KEY:  putative transmembrane region I
;      LOCATION:  350 to 370
;      IDENTIFICATION METHOD:  similarity to other G
;      IDENTIFICATION METHOD:  protein-coupled receptor transmembrane regions,
;      IDENTIFICATION METHOD:  hydrophobic, about 20-23 amino acids in length
;    FEATURE:
;      NAME/KEY:  putative transmembrane region II
;      LOCATION:  382 to 404
;      IDENTIFICATION METHOD:  similarity to other G
;      IDENTIFICATION METHOD:  protein-coupled receptor transmembrane regions,
;      IDENTIFICATION METHOD:  hydrophobic, about 20-23 amino acids in length
;    FEATURE:
;      NAME/KEY:  putative transmembrane region III
;      LOCATION:  427 to 448
;      IDENTIFICATION METHOD:  similarity to other G
;      IDENTIFICATION METHOD:  protein-coupled receptor transmembrane regions,
;      IDENTIFICATION METHOD:  hydrophobic, about 20-23 amino acids in length
;    FEATURE:
;      NAME/KEY:  putative transmembrane region IV
;      LOCATION:  469 to 491
;      IDENTIFICATION METHOD:  similarity to other G
;      IDENTIFICATION METHOD:  protein-coupled receptor transmembrane regions,
;      IDENTIFICATION METHOD:  hydrophobic, about 20-23 amino acids in length
;    FEATURE:
;      NAME/KEY:  putative transmembrane region V
;      LOCATION:  512 to 533
;      IDENTIFICATION METHOD:  similarity to other G
;      IDENTIFICATION METHOD:  protein-coupled receptor transmembrane regions,
;      IDENTIFICATION METHOD:  hydrophobic, about 20-23 amino acids in length
;    FEATURE:
;      NAME/KEY:  putative transmembrane region VI
;      LOCATION:  557 to 580
;      IDENTIFICATION METHOD:  similarity to other G
;      IDENTIFICATION METHOD:  protein-coupled receptor transmembrane regions,
;      IDENTIFICATION METHOD:  hydrophobic, about 20-23 amino acids in length
;    FEATURE:
;      NAME/KEY:  putative transmembrane region VII
;      LOCATION:  592 to 613
;      IDENTIFICATION METHOD:  similarity to other G
;      IDENTIFICATION METHOD:  protein-coupled receptor transmembrane regions,
;      IDENTIFICATION METHOD:  hydrophobic, about 20-23 amino acids in length
;    FEATURE:
;      NAME/KEY:  putative carboxy-terminal intracellular
;      NAME/KEY:  domain
;      LOCATION:  614 to 678
US-08-487-886-2

  Query Match             22.8%;  Score 125;  DB 1;  Length 695;
  Best Local Similarity   26.4%;  Pred. No. 3.6e-06;
  Matches   28;  Conservative   28;  Mismatches   42;  Indels    8;  Gaps    3;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|:|||::|   || :: :::   : ::       :  :| |   ::|:| :|  |:
Db     537 LNVLAFVVICGCYIHIYLTVRNPNIVSS------SSDTRIAKRMAMLIFTDFLCMAPISF 590

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTL-TTNFFKD 105
             |  :  :|  |  :    ::: | |:||  || || : | || :|
Db     591 FAISASLKVPLITVSKAKILLVLFHPINSCANPFLYAIFTKNFRRD 636


Search completed:  May 22, 2003, 12:48:30
Job time :  30 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 22, 2003, 12:45:05 ; Search time 21 Seconds
(without alignments)
505.308 Million cell updates/sec

Title: US-09-930-312-2
Perfect score: 549
Sequence: 1 GVNLLAFLIIVFSYITMFCS.....NSALNPILYTLTTNFFKDKL 107

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	549	100.0	107	9	US-09-930-312-2	Sequence 2, Appli
2	549	100.0	646	10	US-09-928-175-13	Sequence 13, Appl
3	549	100.0	682	10	US-09-928-175-12	Sequence 12, Appl
4	549	100.0	694	10	US-09-928-175-8	Sequence 8, Appli
5	549	100.0	713	9	US-09-965-536A-6	Sequence 6, Appli
6	549	100.0	718	10	US-09-928-175-3	Sequence 3, Appli
7	549	100.0	730	10	US-09-928-175-7	Sequence 7, Appli
8	549	100.0	737	9	US-09-965-536A-2	Sequence 2, Appli
9	549	100.0	754	9	US-10-229-735-1	Sequence 1, Appli
10	549	100.0	754	9	US-10-222-668-2	Sequence 2, Appli
11	549	100.0	754	10	US-09-928-175-2	Sequence 2, Appli
12	478	87.1	718	10	US-09-928-175-21	Sequence 21, Appl
13	478	87.1	737	9	US-10-229-735-3	Sequence 3, Appli
14	478	87.1	737	10	US-09-928-175-20	Sequence 20, Appl
15	376	68.5	176	9	US-09-989-442-105	Sequence 105, App
16	376	68.5	176	9	US-10-073-865-84	Sequence 84, Appl
17	376	68.5	176	9	US-10-103-313-373	Sequence 373, App
18	376	68.5	176	10	US-09-764-853-557	Sequence 557, App
19	376	68.5	176	10	US-09-764-877-2009	Sequence 2009, Ap
20	376	68.5	396	10	US-09-895-686-4	Sequence 4, Appli
21	376	68.5	757	10	US-09-928-175-24	Sequence 24, Appl
22	371	67.6	327	9	US-09-782-974C-66	Sequence 66, Appl
23	234	42.6	334	9	US-10-270-333-21	Sequence 21, Appl
24	234	42.6	722	9	US-10-222-668-4	Sequence 4, Appli
25	225	41.0	359	9	US-10-270-333-27	Sequence 27, Appl
26	210	38.3	1115	9	US-09-965-536A-10	Sequence 10, Appl
27	210	38.3	1115	9	US-10-222-668-3	Sequence 3, Appli
28	142	25.9	861	10	US-09-804-551B-20	Sequence 20, Appl
29	138.5	25.2	458	10	US-09-862-767A-5	Sequence 5, Appli
30	138.5	25.2	699	10	US-09-804-626-6	Sequence 6, Appli
31	136.5	24.9	674	10	US-09-877-804-3	Sequence 3, Appli
32	136.5	24.9	700	10	US-09-877-804-2	Sequence 2, Appli
33	132.5	24.1	693	9	US-09-965-536A-14	Sequence 14, Appl
34	129.5	23.6	676	9	US-09-965-536A-15	Sequence 15, Appl
35	125	22.8	458	10	US-09-862-767A-9	Sequence 9, Appli
36	125	22.8	695	10	US-09-804-626-8	Sequence 8, Appli
37	124	22.6	25	9	US-09-965-536A-21	Sequence 21, Appl
38	123	22.4	326	9	US-09-800-274-4	Sequence 4, Appli
39	123	22.4	915	9	US-10-270-336-5	Sequence 5, Appli
40	122	22.2	675	10	US-09-877-804-7	Sequence 7, Appli
41	122	22.2	687	9	US-09-965-536A-13	Sequence 13, Appl
42	122	22.2	688	9	US-09-965-536A-12	Sequence 12, Appl
43	122	22.2	692	9	US-09-965-536A-11	Sequence 11, Appl
44	122	22.2	692	10	US-09-877-804-6	Sequence 6, Appli
45	120	21.9	794	9	US-10-270-336-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-09-930-312-2
; Sequence 2, Application US/09930312
; Publication No. US20030032019A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn Company
; APPLICANT: Lind, Peter
; TITLE OF INVENTION: No. US20030032019A1el G Protein-Coupled Receptors
; FILE REFERENCE: 00329.PCT1
; CURRENT APPLICATION NUMBER: US/09/930,312
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: USSN 60/225,262
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-930-312-2

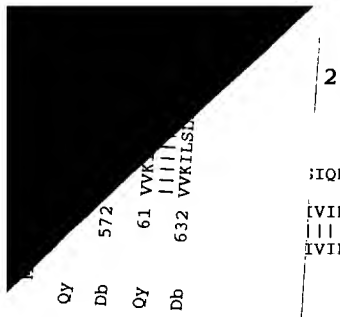
Query Match 100.0%; Score 549; DB 9; Length 107;
Best Local Similarity 100.0%; Pred. No. 3.9e-56;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60

Qy      61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
           |||||||||||||||||||||||||||||||||||||||||||||||
Db      61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
```

RESULT 2
US-09-928-175-13
; Sequence 13, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher J.
; APPLICANT: Gong, Jianhua
; APPLICANT: Daugherty, Betsy

```
;  APPLICANT:  Rogers, No. US20020123618A1ma
;  TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
;  TITLE OF INVENTION:  Uses Thereof
;  FILE REFERENCE: 00-1229
;  CURRENT APPLICATION NUMBER: US/09/928,175
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: 60/224,455
;  PRIOR FILING DATE: 2000-08-10
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
;   LENGTH: 646
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-928-175-13

  Query Match             100.0%;  Score 549;  DB 10;  Length 646;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-55;
  Matches  107;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 547

Qy     61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    548 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 594


RESULT 3
US-09-928-175-12
; Sequence 12, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
;  APPLICANT: Paszty, Christopher J.
;  APPLICANT:  Gong, Jianhua
;  APPLICANT:  Daugherty, Betsy
;  APPLICANT:  Rogers, No. US20020123618A1ma
;  TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
;  TITLE OF INVENTION:  Uses Thereof
;  FILE REFERENCE: 00-1229
;  CURRENT APPLICATION NUMBER: US/09/928,175
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: 60/224,455
;  PRIOR FILING DATE: 2000-08-10
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
;   LENGTH: 682
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-928-175-12

  Query Match             100.0%;  Score 549;  DB 10;  Length 682;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-55;
  Matches  107;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    524 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 583

Qy     61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    584 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 630


RESULT 4
US-09-928-175-8
; Sequence 8, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
;  APPLICANT: Paszty, Christopher J.
;  APPLICANT:  Gong, Jianhua
;  APPLICANT:  Daugherty, Betsy
;  APPLICANT:  Rogers, No. US20020123618A1ma
;  TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
;  TITLE OF INVENTION:  Uses Thereof
;  FILE REFERENCE: 00-1229
;  CURRENT APPLICATION NUMBER: US/09/928,175
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: 60/224,455
;  PRIOR FILING DATE: 2000-08-10
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
;   LENGTH: 694
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-928-175-8

  Query Match             100.0%;  Score 549;  DB 10;  Length 694;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-55;
  Matches  107;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    536 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 595

Qy     61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    596 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 642


RESULT 5
US-09-965-536A-6
; Sequence 6, Application US/09965536A
; Publication No. US20030027323A1
; GENERAL INFORMATION:
;  APPLICANT: FEDER, J. N.
;  APPLICANT:  MINTIER, G.
;  APPLICANT:  RAMANATHAN, C. S.
;  APPLICANT:  HAWKEN, D. R.
;  TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMY5,
;  TITLE OF INVENTION:  EXPRESSED HIGHLY IN BRAIN AND OVARIAN TISSUES
;  FILE REFERENCE: DOO41NP
;  CURRENT APPLICATION NUMBER: US/09/965,536A
;  CURRENT FILING DATE:  2001-09-26
;  PRIOR APPLICATION NUMBER: 60/235,713
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: 60/261,781
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: 60/306,605
;  PRIOR FILING DATE: 2001-07-19
;  PRIOR APPLICATION NUMBER: 60/310,436
;  PRIOR FILING DATE: 2001-08-03
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 713
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-965-536A-6

  Query Match             100.0%;  Score 549;  DB 9;  Length 713;
  Best Local Similarity   100.0%;  Pred. No. 3e-55;
  Matches  107;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    555 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 614

Qy     61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    615 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 661
```

```
RESULT 6
US-09-928-175-3
; Sequence 3, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
;  APPLICANT: Paszty, Christopher J.
;  APPLICANT:  Gong, Jianhua
;  APPLICANT:  Daugherty, Betsy
;  APPLICANT:  Rogers, No. US20020123618A1ma
;  TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
;  TITLE OF INVENTION:  Uses Thereof
;  FILE REFERENCE: 00-1229
;  CURRENT APPLICATION NUMBER: US/09/928,175
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: 60/224,455
;  PRIOR FILING DATE: 2000-08-10
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 718
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-928-175-3

  Query Match           100.0%;  Score 549;  DB 10;  Length 718;
  Best Local Similarity  100.0%;  Pred. No. 3e-55;
  Matches  107;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    560 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 619

Qy     61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    620 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 666


RESULT 7
US-09-928-175-7
; Sequence 7, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
;  APPLICANT: Paszty, Christopher J.
;  APPLICANT:  Gong, Jianhua
;  APPLICANT:  Daugherty, Betsy
;  APPLICANT:  Rogers, No. US20020123618A1ma
;  TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
;  TITLE OF INVENTION:  Uses Thereof
;  FILE REFERENCE: 00-1229
;  CURRENT APPLICATION NUMBER: US/09/928,175
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: 60/224,455
;  PRIOR FILING DATE: 2000-08-10
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 730
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-928-175-7

  Query Match           100.0%;  Score 549;  DB 10;  Length 730;
  Best Local Similarity  100.0%;  Pred. No. 3.1e-55;
  Matches  107;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    631 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 690

Qy     61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    691 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 737


RESULT 8
US-09-965-536A-2
; Sequence 2, Application US/09965536A
; Publication No. US20030027323A1
; GENERAL INFORMATION:
;  APPLICANT: FEDER, J. N.
;  APPLICANT:  MINTIER, G.
;  APPLICANT:  RAMANATHAN, C. S.
;  APPLICANT:  HAWKEN, D. R.
;  TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMY5,
;  TITLE OF INVENTION:  EXPRESSED HIGHLY IN BRAIN AND OVARIAN TISSUES
;  FILE REFERENCE: DO041NP
;  CURRENT APPLICATION NUMBER: US/09/965,536A
;  CURRENT FILING DATE:  2001-09-26
;  PRIOR APPLICATION NUMBER: 60/235,713
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: 60/261,781
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: 60/306,605
;  PRIOR FILING DATE: 2001-07-19
;  PRIOR APPLICATION NUMBER: 60/310,436
;  PRIOR FILING DATE: 2001-08-03
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 737
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-965-536A-2

  Query Match           100.0%;  Score 549;  DB 9;  Length 737;
  Best Local Similarity  100.0%;  Pred. No. 3.1e-55;
  Matches  107;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    579 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 638

Qy     61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    639 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 685


RESULT 9
US-10-229-735-1
; Sequence 1, Application US/10229735
; Publication No. US20030082650A1
; GENERAL INFORMATION:
;  APPLICANT: Baylor College of Medicine
;  APPLICANT:  Agoulnik, Alexander I.
;  TITLE OF INVENTION: The GREAT Gene and Protein
;  FILE REFERENCE: 7572/73263
;  CURRENT APPLICATION NUMBER: US/10/229,735
;  CURRENT FILING DATE:  2002-08-29
;  PRIOR APPLICATION NUMBER: 60/315,696
;  PRIOR FILING DATE: 2001-08-30
;  PRIOR APPLICATION NUMBER: 60/351,432
;  PRIOR FILING DATE: 2002-01-28
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 754
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-229-735-1

  Query Match           100.0%;  Score 549;  DB 9;  Length 754;
  Best Local Similarity  100.0%;  Pred. No. 3.2e-55;
  Matches  107;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```



Qy 572 IQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 655
Db 61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
Qy 632 IVIFFLPVNSALNPILYTLTTNFFKDKL 702
Db

10222668
84A1

, APPLICANT: Hsu, Sheau
, APPLICANT: Hsueh, Aaron
, TITLE OF INVENTION: Mammalian Relaxin Receptor
, FILE REFERENCE: STAN-239 WO
, CURRENT APPLICATION NUMBER: US/10/222,668
, CURRENT FILING DATE: 2002-08-15
, PRIOR APPLICATION NUMBER: 60/313,259
, PRIOR FILING DATE: 2002-08-17
, NUMBER OF SEQ ID NOS: 4
, SOFTWARE: FastSEQ for Windows Version 4.0
, SEQ ID NO 2
, LENGTH: 754
, TYPE: PRT
, ORGANISM: H. sapiens
US-10-222-668-2

Query Match 100.0%; Score 549; DB 9; Length 754;
Best Local Similarity 100.0%; Pred. No. 3.2e-55;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
Db 596 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 655

Qy 61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
Db 656 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 702

RESULT 11
US-09-928-175-2
, Sequence 2, Application US/09928175
, Patent No. US20020123618A1
, GENERAL INFORMATION:
, APPLICANT: Paszty, Christopher J.
, APPLICANT: Gong, Jianhua
, APPLICANT: Daugherty, Betsy
, APPLICANT: Rogers, No. US20020123618A1ma
, TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
, TITLE OF INVENTION: Uses Thereof
, FILE REFERENCE: 00-1229
, CURRENT APPLICATION NUMBER: US/09/928,175
, CURRENT FILING DATE: 2001-08-10
, PRIOR APPLICATION NUMBER: 60/224,455
, PRIOR FILING DATE: 2000-08-10
, NUMBER OF SEQ ID NOS: 42
, SOFTWARE: PatentIn Ver. 2.0
, SEQ ID NO 2
, LENGTH: 754
, TYPE: PRT
, ORGANISM: Homo sapiens
US-09-928-175-2

Query Match 100.0%; Score 549; DB 10; Length 754;
Best Local Similarity 100.0%; Pred. No. 3.2e-55;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
Db 596 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 655

Qy 61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
Db 656 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 702

RESULT 12
US-09-928-175-21
; Sequence 21, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher J.
; APPLICANT: Gong, Jianhua
; APPLICANT: Daugherty, Betsy
; APPLICANT: Rogers, No. US20020123618A1ma
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 00-1229
; CURRENT APPLICATION NUMBER: US/09/928,175
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/224,455
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 718
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-928-175-21

Query Match 87.1%; Score 478; DB 10; Length 718;
Best Local Similarity 86.0%; Pred. No. 4.7e-47;
Matches 92; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

Qy 1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
Db 560 GVNLLAFLVIVISYVTMFCSIHKTALQTAEVRSHIGKEVAVANRFFFIVFSDAICWIPVF 619

Qy 61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
Db 620 VVKILSLLQVEIPGTITSWIVVFFLPVNSALNPILYTLTTSFFKDKL 666

RESULT 13
US-10-229-735-3
; Sequence 3, Application US/10229735
; Publication No. US20030082650A1
; GENERAL INFORMATION:
; APPLICANT: Baylor College of Medicine
; APPLICANT: Agoulnik, Alexander I.
; TITLE OF INVENTION: The GREAT Gene and Protein
; FILE REFERENCE: 7572/73263
; CURRENT APPLICATION NUMBER: US/10/229,735
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/315,696
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 60/351,432
; PRIOR FILING DATE: 2002-01-28
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 737
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-229-735-3

Query Match 87.1%; Score 478; DB 9; Length 737;
Best Local Similarity 86.0%; Pred. No. 4.9e-47;
Matches 92; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

Qy 1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
Db 579 GVNLLAFLVIVISYVTMFCSIHKTALQTAEVRSHIGKEVAVANRFFFIVFSDAICWIPVF 638

```
Qy      61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
           ||||||| :|||| |:|||||:||||||||||||||||||:||||||
Db     639 VVKILSLLQVEIPGTITSWIVVFFLPVNSALNPILYTLTTSFFKDKL 685


RESULT 14
US-09-928-175-20
; Sequence 20, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
;  APPLICANT: Paszty, Christopher J.
;  APPLICANT:  Gong, Jianhua
;  APPLICANT:  Daugherty, Betsy
;  APPLICANT:  Rogers, No. US20020123618A1ma
;  TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
;  TITLE OF INVENTION:  Uses Thereof
;  FILE REFERENCE: 00-1229
;  CURRENT APPLICATION NUMBER: US/09/928,175
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: 60/224,455
;  PRIOR FILING DATE: 2000-08-10
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
;   LENGTH: 737
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-928-175-20

  Query Match             87.1%;  Score 478;  DB 10;  Length 737;
  Best Local Similarity   86.0%;  Pred. No. 4.9e-47;
  Matches   92;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
           |||||||:|| ||:|||||| |||||| |||:  |:||||||||||||||||||||||
Db     579 GVNLLAFLVIVISYVTMFCSIHKTALQTAEVRSHIGKEVAVANRFFFIVFSDAICWIPVF 638

Qy      61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
           ||||||| :|||| |:|||||:||||||||||||||||||:||||||
Db     639 VVKILSLLQVEIPGTITSWIVVFFLPVNSALNPILYTLTTSFFKDKL 685


RESULT 15
US-09-989-442-105
; Sequence 105, Application US/09989442
; Publication No. US20030013649A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PJZ08
;  CURRENT APPLICATION NUMBER: US/09/989,442
;  CURRENT FILING DATE:  2001-11-21
;  PRIOR APPLICATION NUMBER: 60/179,065
;  PRIOR FILING DATE: 2000-01-31
;  PRIOR APPLICATION NUMBER: 60/180,628
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: 60/214,886
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/217,487
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,758
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/220,963
;  PRIOR FILING DATE: 2000-07-26
;  PRIOR APPLICATION NUMBER: 60/217,496
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,447
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/218,290
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/225,757
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/226,868
;  PRIOR FILING DATE: 2000-08-22
;  PRIOR APPLICATION NUMBER: 60/216,647
;  PRIOR FILING DATE: 2000-07-07
;  PRIOR APPLICATION NUMBER: 60/225,267
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/216,880
;  PRIOR FILING DATE: 2000-07-07
;  PRIOR APPLICATION NUMBER: 60/225,270
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/251,869
;  PRIOR FILING DATE: 2000-12-08
;  PRIOR APPLICATION NUMBER: 60/235,834
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: 60/234,274
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: 60/234,223
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: 60/228,924
;  PRIOR FILING DATE: 2000-08-30
;  PRIOR APPLICATION NUMBER: 60/224,518
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/236,369
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/224,519
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/220,964
;  PRIOR FILING DATE: 2000-07-26
;  PRIOR APPLICATION NUMBER: 60/241,809
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/249,299
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/236,327
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/241,785
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/244,617
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 60/225,268
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/236,368
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/251,856
;  PRIOR FILING DATE: 2000-12-08
;  PRIOR APPLICATION NUMBER: 60/251,868
;  PRIOR FILING DATE: 2000-12-08
;  PRIOR APPLICATION NUMBER: 60/229,344
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 60/234,997
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: 60/229,343
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 60/229,345
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 60/229,287
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 60/229,513
;  PRIOR FILING DATE: 2000-09-05
;  PRIOR APPLICATION NUMBER: 60/231,413
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/229,509
;  PRIOR FILING DATE: 2000-09-05
;  PRIOR APPLICATION NUMBER: 60/236,367
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/237,039
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: 60/237,038
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: 60/236,370
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/236,802
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: 60/237,037
;  PRIOR FILING DATE: 2000-10-02
```

; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,065
; PRIOR FILING DATE: 2000-09-14

Query Match 68.5%; Score 376; DB 9; Length 176;
Best Local Similarity 64.5%; Pred. No. 6.1e-36;
Matches 69; Conservative 20; Mismatches 18; Indels 0; Gaps 0;

```
Qy       1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
           |:|| ||:|||||| :|| |: ::|:  ||:||   :|: :| |||||||:||:||||:|
Db      11 GINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVFTDALCWIPIF 70

Qy      61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
             ||| ||| :|||| |:|||:||| ||:||||||||||||||  ||: :
Db      71 VVKFLSLLQVEIPGTITSWVVIFILPINSALNPILYTLTTRPFKEMI 117
```

Search completed: May 22, 2003, 12:48:58
Job time : 22 secs

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 22, 2003, 12:43:05 ; Search time 43 Seconds
(without alignments)
239.218 Million cell updates/sec

Title: US-09-930-312-2
Perfect score: 549
Sequence: 1 GVNLLAFLIIVFSYITMFCS.....NSALNPILYTLTTNFFKDKL 107

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	210	38.3	1115	2	S40241	G protein-coupled
2	145.5	26.5	320	2	S17177	probable G protein
3	145.5	26.5	337	2	S68678	adenosine receptor
4	139.5	25.4	696	2	A41344	lutropin-choriogon
5	139	25.3	925	2	JC2033	G protein-coupled
6	138.5	25.2	320	2	A46152	A3 adenosine recep
7	138.5	25.2	699	1	QRHUUT	lutropin-choriogon
8	136.5	24.9	700	2	I77463	luteinizing hormon
9	136.5	24.9	700	2	A49744	lutropin-choriogon
10	135.5	24.7	814	2	JC7389	thyroid stimulatin
11	134.5	24.5	793	2	JC7390	thyroid stimulatin
12	132.5	24.1	700	2	A42395	lutropin receptor
13	132	24.0	694	2	JC4301	follitropin recept
14	129	23.5	694	2	JC2237	follitropin recept
15	125.5	22.9	696	2	JC7361	follitropin recept
16	125	22.8	695	1	QRHUFT	follitropin recept
17	125	22.8	695	2	I45896	follicle stimulati
18	125	22.8	695	2	JC1493	follitropin recept
19	124	22.6	326	2	I48096	A1 adenosine recep
20	123	22.4	326	2	A53005	adenosine receptor
21	122	22.2	692	2	A34548	follitropin recept
22	122	22.2	695	1	JN0898	follitropin recept
23	121	22.0	324	2	S55302	A(1) adenosine rec
24	120.5	21.9	764	2	A40077	thyrotropin recept
25	120.5	21.9	764	2	JC5643	thyroid stimulatin
26	120	21.9	185	2	I48931	adenosine receptor
27	120	21.9	326	2	A40376	adenosine receptor
28	119.5	21.8	318	2	S38511	adenosine receptor
29	119.5	21.8	764	1	QRHURH	thyrotropin recept
30	119.5	21.8	764	2	A35956	thyrotropin recept
31	119	21.7	326	2	C30341	G protein-coupled
32	119	21.7	332	2	I48933	adenosine receptor
33	119	21.7	907	2	JG0193	G protein-coupled
34	118.5	21.6	764	2	I48882	thyrotropin recept
35	117	21.3	326	2	A38144	adenosine receptor
36	116	21.1	332	2	A42171	A2-adenosine recep
37	115	20.9	332	2	JC1229	adenosine receptor
38	114	20.8	328	2	JN0675	adenosine receptor
39	110	20.0	907	2	JE0176	orphan G protein-c
40	109.5	19.9	328	2	T30999	hypothetical prote
41	108	19.7	391	2	A41795	somatostatin recep
42	108	19.7	391	2	C41795	somatostatin recep
43	108	19.7	391	2	A39297	somatostatin recep
44	107	19.5	386	2	S72168	dopamine receptor
45	105.5	19.2	381	2	A35300	G protein-coupled

ALIGNMENTS

RESULT 1
S40241
G protein-coupled receptor - great pond snail
C;Species: Lymnaea stagnalis (great pond snail)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 15-Jun-2001
C;Accession: S40241
R;Tensen, C.P.; Kesteren, E.R.; Planta, R.J.; Cox, K.; Burke, J.F.; Heerikhuizen, H.; submitted to the EMBL Data Library, June 1993
A;Description: A G protein-coupled receptor with LDL-binding motifs suggests a role f
A;Reference number: S40241
A;Accession: S40241
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1115 <TEN>
A;Cross-references: EMBL:Z23104; NID:g438128; PID:g438129
C;Superfamily: great pond snail LDL receptor-related G protein-coupled receptor; LDL
C;Keywords: G protein-coupled receptor; transmembrane protein
F;38-77/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F;79-113/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F;118-153/Domain: LDL receptor ligand-binding repeat homology <LDLA>
F;158-194/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F;195-230/Domain: LDL receptor ligand-binding repeat homology <LDLB>
F;233-267/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F;274-316/Domain: LDL receptor ligand-binding repeat homology <LDLC>
F;322-361/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F;367-401/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F;406-440/Domain: LDL receptor ligand-binding repeat homology <LDLD>
F;446-483/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F;488-523/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F;584-607/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;608-631/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;632-655/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;656-679/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;704-727/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;774-797/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>

Query Match 38.3%; Score 210; DB 2; Length 1115;
Best Local Similarity 41.6%; Pred. No. 4.2e-15;
Matches 42; Conservative 23; Mismatches 30; Indels 6; Gaps 2;

```
Qy       2 VNLLAFLIIVFSYITMFCSIQKT--ALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPV 59
           :|||:|::|  ||: ||   :||  |::| | :|    : |:| |   || :|  ||:|:
Db     946 LNLLSFVLIASSYLWMFSVAKKTRSAVRTAESKN----DNAMARRMTLIVMTDFCCWVPI 1001

Qy      60 FVVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTT 100
            |:  :||      | : :|| :| ||:||| ||::||||:|
Db    1002 IVLGFVSLAGARADDQVYAWIAVFVLPLNSATNPVIYTLST 1042
```

RESULT 2
S17177
probable G protein-coupled receptor - rat

```
C;Species: Rattus norvegicus (Norway rat)
C;Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 29-Sep-1999
C;Accession: S17177
R;Meyerhof, W.; Mueller-Brechlin, R.; Richter, D.
FEBS Lett. 284, 155-160, 1991
A;Title: Molecular cloning of a novel putative G-protein coupled receptor expressed duri
A;Reference number: S17177; MUID:91285122; PMID:1647979
A;Accession: S17177
A;Molecule type: mRNA
A;Residues: 1-320 <MEY>
A;Cross-references: EMBL:X59249; NID:g56307; PIDN:CAA41937.1; PID:g56308
C;Superfamily: adenosine receptor A1
C;Keywords: G protein-coupled receptor; glycoprotein; lipoprotein; thiolester bond; tran
F;17-39/Domain: transmembrane #status predicted <TM1>
F;51-74/Domain: transmembrane #status predicted <TM2>
F;84-108/Domain: transmembrane #status predicted <TM3>
F;129-150/Domain: transmembrane #status predicted <TM4>
F;180-200/Domain: transmembrane #status predicted <TM5>
F;234-257/Domain: transmembrane #status predicted <TM6>
F;264-286/Domain: transmembrane #status predicted <TM7>
F;4,5,42,162,213/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;302/Binding site: palmitate (Cys) (covalent) #status predicted

  Query Match             26.5%;  Score 145.5;  DB 2;  Length 320;
  Best Local Similarity   29.2%;  Pred. No. 1.6e-08;
  Matches   31;  Conservative   23;  Mismatches   47;  Indels    5;  Gaps    2;

Qy       4 LLAFLIIVFSYITMFCSIQKTALQT----TEVRNCFGREVAVANRFFFIVFSDAICWIPV 59
           |:  :::   |: :|  |:    |       | |  :|||   |   | ::|  |:||:|:
Db     189 LIPLVVMCIIYLDIFYIIRNKLSQNLTGFRETRAFYGREFKTAKSLFLVLFLFALCWLPL 248

Qy      60 FVVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKD 105
             ::  :| | |:||: :   : |     || :|||:|      ||:
Db     249 SIINFVSYFNVKIPE-IAMCLGILLSHANSMMNPIVYACKIKKFKE 293


RESULT 3
S68678
adenosine receptor A3i - rat
C;Species: Rattus sp. (rat)
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C;Accession: S68678
R;Sajjadi, F.G.; Boyle, D.L.; Domingo, R.C.; Firestein, G.S.
FEBS Lett. 382, 125-129, 1996
A;Title: cDNA cloning and characterization of A3i, an alternatively spliced rat A3 adenc
A;Reference number: S68678; MUID:96196578; PMID:8612733
A;Accession: S68678
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-337 <SAJ>
A;Cross-references: EMBL:X93219; NID:g1297000; PIDN:CAA63702.1; PID:g1297001
C;Superfamily: adenosine receptor A1
C;Keywords: alternative splicing

  Query Match             26.5%;  Score 145.5;  DB 2;  Length 337;
  Best Local Similarity   29.2%;  Pred. No. 1.7e-08;
  Matches   31;  Conservative   23;  Mismatches   47;  Indels    5;  Gaps    2;

Qy       4 LLAFLIIVFSYITMFCSIQKTALQT----TEVRNCFGREVAVANRFFFIVFSDAICWIPV 59
           |:  :::   |: :|  |:    |       | |  :|||   |   | ::|  |:||:|:
Db     206 LIPLVVMCIIYLDIFYIIRNKLSQNLTGFRETRAFYGREFKTAKSLFLVLFLFALCWLPL 265

Qy      60 FVVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKD 105
             ::  :| | |:||: :   : |     || :|||:|      ||:
Db     266 SIINFVSYFNVKIPE-IAMCLGILLSHANSMMNPIVYACKIKKFKE 310


RESULT 4
A41344
lutropin-choriogonadotropin receptor precursor (splice form A) - pig
N;Alternate names: luteinizing hormone-choriogonadotropin receptor
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 21-Jan-2000
C;Accession: A41344
R;Loosfelt, H.; Misrahi, M.; Atger, M.; Salesse, R.; Vu Hai-Luu Thi, M.T.; Jolivet, A
Science 245, 525-528, 1989
A;Title: Cloning and sequencing of porcine LH-hCG receptor cDNA: variants lacking tra
A;Reference number: A41344; MUID:89332517; PMID:2502844
A;Accession: A41344
A;Molecule type: mRNA
A;Residues: 1-696 <LOO>
A;Cross-references: GB:M29525; NID:g164528; PIDN:AAA31062.1; PID:g164529
C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repea
C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; membrane
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-696/Product: lutropin-choriogonadotropic hormone receptor #status predicted <MAT
F;99-123/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

  Query Match             25.4%;  Score 139.5;  DB 2;  Length 696;
  Best Local Similarity   28.8%;  Pred. No. 1.6e-07;
  Matches   30;  Conservative   24;  Mismatches   43;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|::||:||   || :: ::|   |  |       ::  :| :   ::|:|  |  |:
Db     534 LNVVAFIIICACYIKIYFAVQNPELMAT------NKDTKIAKKMAVLIFTDFTCMAPISF 587

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
             | :  :|  |  | :  ::: | ||||  || || : |  |:
Db     588 FAISAALKVPLITVTNSKVLLVLFYPVNSCANPFLYAIFTKAFR 631


RESULT 5
JC2033
G protein-coupled receptor - sea anemone (Anthopleura elegantissima)
C;Species: Anthopleura elegantissima
C;Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 21-Jul-2000
C;Accession: JC2033; S41908
R;Nothacker, H.P.; Grimmelikhuijzen, C.J.P.
Biochem. Biophys. Res. Commun. 197, 1062-1069, 1993
A;Title: Molecular cloning of a novel, putative G protein-coupled receptor from sea a
A;Reference number: JC2033; MUID:94107299; PMID:8280121
A;Accession: JC2033
A;Molecule type: mRNA
A;Residues: 1-925 <NOT>
A;Cross-references: EMBL:Z28332; NID:g452810; PIDN:CAA82186.1; PID:g452811
C;Superfamily: leucine-rich alpha-2-glycoprotein repeat homology
C;Keywords: G protein-coupled receptor; glycoprotein; neuropeptide; phosphoprotein; t
F;61-84/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;85-109/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;110-133/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;134-157/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;333-349,350-384,385-419,420-453,454-462/Region: glycine-rich
F;530-551/Domain: transmembrane #status predicted <TM1>
F;562-584/Domain: transmembrane #status predicted <TM2>
F;607-628/Domain: transmembrane #status predicted <TM3>
F;652-673/Domain: transmembrane #status predicted <TM4>
F;692-712/Domain: transmembrane #status predicted <TM5>
F;740-763/Domain: transmembrane #status predicted <TM6>
F;775-795/Domain: transmembrane #status predicted <TM7>
F;61,152,212,435,491,495/Binding site: carbohydrate (Asn) (covalent) #status predicte
F;553,645,855,859,868,894,900/Binding site: phosphate (Ser) (covalent) #status predic
F;649,853/Binding site: phosphate (Thr) (covalent) #status predicted

  Query Match             25.3%;  Score 139;  DB 2;  Length 925;
  Best Local Similarity   29.6%;  Pred. No. 2.4e-07;
  Matches   32;  Conservative   24;  Mismatches   46;  Indels    6;  Gaps    3;

Qy       2 VNLLAFLIIVFSYITM-FCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
           ||  :|: ::: |  | :  :    .::    ||    :  || |   :||:| :|| |:
Db     700 VNGASFISVMYLYSRMLYVVVSGGDMEGAPKRN----DSKVAKRMAILVFTDMLCWAPIA 755

Qy      61 VVKILSLF-RVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
             :|: | :  :  | :  :::|| |:||  || ||   |  || :|
Db     756 FFGLLAAFGQTLLTVTQSKILLVFFFPINSICNPFLYAFFTKAFKREL 803
```

```
RESULT 6
A46152
A3 adenosine receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 29-Sep-1999
C;Accession: A46152
R;Zhou, Q.Y.; Li, C.; Olah, M.E.; Johnson, R.A.; Stiles, G.L.; Civelli, O.
Proc. Natl. Acad. Sci. U.S.A. 89, 7432-7436, 1992
A;Title: Molecular cloning and characterization of an adenosine receptor: the A3 adenosi
A;Reference number: A46152; MUID:92366475; PMID:1323836
A;Accession: A46152
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-320 <ZHO>
A;Cross-references: GB:M94152; NID:g202724; PIDN:AAA40680.1; PID:g460332
A;Experimental source: brain
A;Note: sequence inconsistent with the nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:110661, NCBIP:110662)
C;Superfamily: adenosine receptor A1
C;Keywords: G protein-coupled receptor; transmembrane protein

  Query Match             25.2%;  Score 138.5;  DB 2;  Length 320;
  Best Local Similarity   29.9%;  Pred. No. 9.4e-08;
  Matches   29;  Conservative   22;  Mismatches   41;  Indels    5;  Gaps    2;

Qy       4 LLAFLIIVFSYITMFCSIQKTALQT----TEVRNCFGREVAVANRFFFIVFSDAICWIPV 59
           |:  :::   |: :|  |:    |      | |  :|||   |   | ::|  |:||:|:
Db     189 LIPLVVMCIIYLDIFYIIRNKLSQNLTGFRETRAFYGREFKTAKSLFLVLFLFALCWLPL 248

Qy      60 FVVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILY 96
            ::  :| | |:||: :   : |     || :|||:|
Db     249 SIINFVSYFNVKIPE-IAMCLGILLSHANSMMNPIVY 284


RESULT 7
QRHUUT
lutropin-choriogonadotropin receptor precursor - human
N;Alternate names: luteinizing hormone-choriogonadotropin receptor
N;Contains: lutropin-choriogonadotropin receptor precursor long splice form; lutropin-ch
C;Species: Homo sapiens (man)
C;Date: 22-Jan-1993 #sequence_revision 16-Feb-1996 #text_change 22-Jun-1999
C;Accession: A36243; B36243; A23728; B36120
R;Minegish, T.; Nakamura, K.; Takakura, Y.; Miyamoto, K.; Hasegawa, Y.; Ibuki, Y.; Igara
Biochem. Biophys. Res. Commun. 172, 1049-1054, 1990
A;Title: Cloning and sequencing of human LH/hCG receptor cDNA.
A;Reference number: A36243; MUID:91058534; PMID:2244890
A;Accession: A36243
A;Molecule type: mRNA
A;Residues: 1-699 <MIN>
A;Cross-references: GB:M63108; NID:g187135; PIDN:AAA59515.1; PID:g187136
A;Experimental source: ovary
A;Accession: B36243
A;Molecule type: mRNA
A;Residues: 1-226,290-699 <MI2>
R;Jia, X.C.; Oikawa, M.; Bo, M.; Tanaka, T.; Ny, T.; Boime, I.; Hsueh, A.J.W.
Mol. Endocrinol. 5, 759-768, 1991
A;Title: Expression of human luteinizing hormone (LH) receptor: interaction with LH and
A;Reference number: A23728; MUID:92017881; PMID:1922095
A;Accession: A23728
A;Molecule type: mRNA
A;Residues: 1-123,'R',125-311,'N',313-699 <JIA>
A;Cross-references: GB:S57793
A;Experimental source: ovary granulosa and luteal cells; thyroid
A;Note: 312-Ser was also found; the authors translated the codon UAC for residue 546 as
A;Note: mRNA from the thyroid is incompletely processed compared with that from the ovar
R;Frazier, A.L.; Robbins, L.S.; Stork, P.J.; Sprengel, R.; Segaloff, D.L.; Cone, R.D.
Mol. Endocrinol. 4, 1264-1276, 1990
A;Title: Isolation of TSH and LH/CG receptor cDNAs from human thyroid: regulation by tis
A;Reference number: A36120; MUID:91155962; PMID:2293030
A;Accession: B36120
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-6,'P',8-18,'A',20-26,'R',29-43,'A',45,'A',47,'RPS',52-67,'S',69-123,'R',12
A;Cross-references: GB:M73746; NID:g903745; PIDN:AAA70231.1; PID:g903746
A;Experimental source: thyroid
R;Zhu, H.; Wang, H.; Ascoli, M.
Mol. Endocrinol. 9, 141-150, 1995
A;Title: The lutropin/choriogonadotropin receptor is palmitoylated at intracellular c
A;Reference number: A57564; MUID:95295727; PMID:7776964
A;Contents: annotation; mutation analysis of palmitate binding sites
C;Genetics:
A;Gene: GDB:LHCGR
A;Cross-references: GDB:125260; OMIM:152790
A;Map position: 2p21-2p21
A;Introns: 226/3
A;Note: the exact position of the intron cannot be determined from the experimental d
C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repea
C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; hormone r
F;1-699/Product: lutropin-choriogonadotropin receptor precursor, long splice form #st
F;1-226,290-699/Product: lutropin-choriogonadotropin receptor precursor, short splice
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-644/Product: lutropin-choriogonadotropin receptor #status predicted <MAT>
F;23-363/Domain: extracellular hormone binding #status predicted <EHB>
F;50-73/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;74-98/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;99-123/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;124-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;149-173/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;176-197/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F;198-222/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F;364-385/Domain: transmembrane #status predicted <TM1>
F;396-418/Domain: transmembrane #status predicted <TM2>
F;441-462/Domain: transmembrane #status predicted <TM3>
F;483-505/Domain: transmembrane #status predicted <TM4>
F;526-547/Domain: transmembrane #status predicted <TM5>
F;571-594/Domain: transmembrane #status predicted <TM6>
F;606-627/Domain: transmembrane #status predicted <TM7>
F;645-699/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F;99,174,195,291,299,313/Binding site: carbohydrate (Asn) (covalent) #status predicte
F;561,677/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status pred
F;643,644/Binding site: palmitate (Cys) (covalent) #status experimental
F;670/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicte

  Query Match             25.2%;  Score 138.5;  DB 1;  Length 699;
  Best Local Similarity   27.9%;  Pred. No. 2e-07;
  Matches   29;  Conservative   25;  Mismatches   43;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
             :|::|| ||   || :: :::   |  |       ::  :| :   ::|:|  |  |:
Db     534 LNVVAFFIICACYIKIYFAVRNPELMAT------NKDTKIAKKMAILIFTDFTCMAPISF 587

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
           | : |:|  |  | :  ::: | |:|| || || : |  |:
Db     588 FAISAAFKVPLITVTNSKVLLVLFYPINSCANPFLYAIFTKTFQ 631


RESULT 8
I77463
luteinizing hormone/chorionic gonadotropin receptor - rat
C;Species: Rattus sp. (rat)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jan-2000
C;Accession: I77463
R;Aatsinki, J.T.; Pietila, E.M.; Lakkakorpi, J.T.; Rajaniemi, H.J.
Mol. Cell. Endocrinol. 84, 127-135, 1992
A;Title: Expression of the LH/CG receptor gene in rat ovarian tissue is regulated by
A;Reference number: I57668; MUID:92347604; PMID:1353463
A;Accession: I77463
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-700 <RES>
A;Cross-references: GB:S40803; NID:g252163; PIDN:AAB22680.1; PID:g252164
C;Genetics:
A;Introns: 58/2; 82/2; 107/2; 132/2; 157/2; 183/2; 206/2; 231/2; 293/2; 320/2
C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repea
F;153-177/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>
```

```
  Query Match             24.9%;  Score 136.5;  DB 2;  Length 700;
  Best Local Similarity   27.9%;  Pred. No. 3.4e-07;
  Matches   29;  Conservative   25;  Mismatches   43;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|:::||::|   || :: ::|   |          ::  :| :   ::|:|  |  |:
Db     538 LNVVAFVVICACYIRIYFAVQNPELTAP------NKDTKIAKKMAILIFTDFTCMAPISF 591

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
             | : |:|  |  | :  ::: | ||||  || || : |  |:   .
Db     592 FAISAAFKVPLITVTNSKILLVLFYPVNSCANPFLYAIFTKAFQ 635


RESULT 9
A49744
lutropin-choriogonadotropin receptor precursor - rat
N;Alternate names: luteinizing hormone-choriogonadotropin receptor
C;Species: Rattus norvegicus (Norway rat)
C;Date: 03-May-1994 #sequence_revision 13-Sep-1998 #text_change 13-Aug-1999
C;Accession: A49744; A40545; A41343; A61453; A32460
R;Koo, Y.B.; Ji, I.; Slaughter, R.G.; Ji, T.H.
Endocrinology 128, 2297-2308, 1991
A;Title: Structure of the luteinizing hormone receptor gene and multiple exons of the cc
A;Reference number: A49744; MUID:91209270; PMID:2019252
A;Accession: A49744
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-700 <KOO>
A;Cross-references: GB:M68917
A;Note: authors translated the codon CAA for residue 307 as Glu, AAC for residue 355 as
R;Tsai-Morris, C.H.; Buczko, E.; Wang, W.; Xie, X.Z.; Dufau, M.L.
J. Biol. Chem. 266, 11355-11359, 1991
A;Title: Structural organization of the rat luteinizing hormone (LH) receptor gene.
A;Reference number: A40545; MUID:91250455; PMID:2040640
A;Accession: A40545
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-320 <TSA>
A;Cross-references: GB:M63918; GB:M63919; GB:M63920; GB:M63921; GB:M63922; GB:M63923; GE
R;McFarland, K.C.; Sprengel, R.; Phillips, H.S.; Koehler, M.; Rosemblit, N.; Nikolics, K
Science 245, 494-499, 1989
A;Title: Lutropin-choriogonadotropin receptor: an unusual member of the G protein-couple
A;Reference number: A41343; MUID:89332512; PMID:2502842
A;Accession: A41343
A;Molecule type: mRNA
A;Residues: 1-700 <MCF>
A;Cross-references: GB:M26199; NID:g205178; PIDN:AAA41528.1; PID:g205179
R;Dufau, M.L.; Minegishi, T.; Buczko, E.S.; Delgado, C.J.; Zhang, R.
J. Steroid Biochem. 33, 715-720, 1989
A;Title: Characterization and structure of ovarian and testicular LH/hCG receptors.
A;Reference number: A61453; MUID:90097014; PMID:2601325
A;Accession: A61453
A;Status: preliminary
A;Molecule type: protein
A;Residues: 27-33,'X',35-37,'X',39,'X',41-44 <DUF>
R;Roche, P.C.; Ryan, R.J.
J. Biol. Chem. 264, 4636-4641, 1989
A;Title: Purification, characterization, and amino-terminal sequence of rat ovarian rece
A;Reference number: A32460; MUID:89174723; PMID:2925659
A;Accession: A32460
A;Molecule type: protein
A;Residues: 27-32,'LX',35-37 <ROC>
C;Genetics:
A;Introns: 58/2; 82/2; 107/2; 132/2; 157/2; 183/2; 206/2; 231/2; 293/2; 320/2
C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; hormone rece
F;54-77/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;78-102/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;103-127/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;128-152/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;153-177/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;180-201/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F;202-226/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F;368-389/Domain: transmembrane #status predicted <TM1>
F;400-422/Domain: transmembrane #status predicted <TM2>
F;444-466/Domain: transmembrane #status predicted <TM3>
F;489-511/Domain: transmembrane #status predicted <TM4>
F;530-551/Domain: transmembrane #status predicted <TM5>
F;575-598/Domain: transmembrane #status predicted <TM6>
F;610-631/Domain: transmembrane #status predicted <TM7>
F;103,178,199,295,303,317/Binding site: carbohydrate (Asn) (covalent) #status predict
F;647,648/Binding site: palmitate (Cys) (covalent) #status predicted
F;681/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicte

  Query Match             24.9%;  Score 136.5;  DB 2;  Length 700;
  Best Local Similarity   27.9%;  Pred. No. 3.4e-07;
  Matches   29;  Conservative   25;  Mismatches   43;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|:::||::|   || :: ::|   |          ::  :| :   ::|:|  |  |:
Db     538 LNVVAFVVICACYIRIYFAVQNPELTAP------NKDTKIAKKMAILIFTDFTCMAPISF 591

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
             | : |:|  |  | :  ::: | ||||  || || : |  |:
Db     592 FAISAAFKVPLITVTNSKILLVLFYPVNSCANPFLYAIFTKAFQ 635


RESULT 10
JC7389
thyroid stimulating hormone receptor a - salmon
N;Alternate names: thyrotropin receptor a
C;Species: Oncorhynchus sp. (salmon)
C;Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 31-Dec-2000
C;Accession: JC7389
R;Oba, Y.; Hirai, T.; Yoshiura, Y.; Kobayashi, T.; Nagahama, Y.
Biochem. Biophys. Res. Commun. 276, 258-263, 2000
A;Title: Cloning, functional characterization, and expression of thyrotropin receptor
A;Reference number: JC7389
A;Contents: Thyroid
A;Accession: JC7389
A;Molecule type: mRNA
A;Residues: 1-814 <OBA>
A;Cross-references: DDBJ:AB030954
C;Comment: This receptor, a transmembrane protein, which mediates the actions of thyr
C;Genetics:
A;Gene: tsh-ra
A;Introns: 610/3
C;Keywords: glycoprotein; hormone receptor; thyroid gland; transmembrane protein

  Query Match             24.7%;  Score 135.5;  DB 2;  Length 814;
  Best Local Similarity   27.9%;  Pred. No. 5.1e-07;
  Matches   29;  Conservative   25;  Mismatches   43;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPV-F 60
           :|:||||:|   || ::|::     ::        ::  :| |   ::|:| :|  |:|
Db     634 LNILAFLVICACYIKIYCTVHNPHYRSG------SKDTNIAKRMAVLIFTDFLCMAPISF 687

Qy      61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
               | |  |  : :  ::: | |:||  || || : |  |:
Db     688 YAMSAVLDRPLITVSNSKILLVLFYPLNSCANPFLYAIFTKAFR 731


RESULT 11
JC7390
thyroid stimulating hormone receptor b - salmon
N;Alternate names: thyrotropin receptor b
C;Species: Oncorhynchus sp. (salmon)
C;Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 31-Dec-2000
C;Accession: JC7390
R;Oba, Y.; Hirai, T.; Yoshiura, Y.; Kobayashi, T.; Nagahama, Y.
Biochem. Biophys. Res. Commun. 276, 258-263, 2000
A;Title: Cloning, functional characterization, and expression of thyrotropin receptor
A;Reference number: JC7389
A;Contents: Thyroid
A;Accession: JC7390
A;Molecule type: mRNA
```

A;Residues: 1-793 <OBA>
A;Cross-references: DDBJ:AB030955
C;Comment: This receptor, a transmembrane protein, which medaites the actions of thyrotr
C;Genetics:
A;Gene: tsh-rb
A;Introns: 584/3
C;Keywords: glycoprotein; hormone receptor; thyroid gland; transmembrane protein

Query Match 24.5%; Score 134.5; DB 2; Length 793;
Best Local Similarity 27.6%; Pred. No. 6.4e-07;
Matches 29; Conservative 25; Mismatches 42; Indels 9; Gaps 3;

Qy 2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFG-REVAVANRFFFIVFSDAICWIPV- 59
Db 608 LNILAFLVICTCYVKIYCAVHNPYY-------CSGSKDTNIAKRMAILIFTDFLCMAPIS 660

Qy 60 FVVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
Db 661 FYAMSAVVDRPLITVSNSKILLVLFYPLNSCANPFLYAIFTKAFR 705

RESULT 12
A42395
lutropin receptor - mouse
N;Alternate names: luteinizing hormone-choriogonadotropin receptor
C;Species: Mus musculus (house mouse)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 13-Aug-1999
C;Accession: A42395
R;Gudermann, T.; Birnbaumer, M.; Birnbaumer, L.
J. Biol. Chem. 267, 4479-4488, 1992
A;Title: Evidence for dual coupling of the murine luteinizing hormone receptor to adenyl
mone receptor expressed in L cells.
A;Reference number: A42395; MUID:92165799; PMID:1311310
A;Accession: A42395
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-700 <GUD>
A;Cross-references: GB:M81310; GB:M81318; NID:g198811; PIDN:AAA39432.1; PID:g198812
A;Note: sequence extracted from NCBI backbone (NCBIN:84064, NCBIP:84066)
C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
C;Keywords: G protein-coupled receptor; transmembrane protein
F;54-77/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;78-102/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;103-127/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;128-152/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;153-177/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;180-201/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F;202-226/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

Query Match 24.1%; Score 132.5; DB 2; Length 700;
Best Local Similarity 26.9%; Pred. No. 9.4e-07;
Matches 28; Conservative 25; Mismatches 44; Indels 7; Gaps 2;

Qy 2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
Db 538 LNAVAFVVICACYVRIYFAVQNPELTAP------NKDTKIAKKMAILIFTDFTCMAPISF 591

Qy 62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
Db 592 FAISAAFKVPLITVTNSKVLLVLFYPVNSCANPFLYAVFTKAFQ 635

RESULT 13
JC4301
follitropin receptor - pig
N;Alternate names: follicle-stimulating hormone receptor
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 16-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 21-Jan-2000
C;Accession: JC4301
R;Remy, J.J.; Lahbib-Mansais, Y.; Yerle, M.; Bozon, V.; Couture, L.; Pajot, E.; Grebert,
Gene 163, 257-261, 1995
A;Title: The porcine follitropin receptor: cDNA cloning, functional expressionand chromo
A;Reference number: JC4301; MUID:96011644; PMID:7590277
A;Accession: JC4301
A;Molecule type: mRNA
A;Residues: 1-694 <REM>
A;Cross-references: GB:L31966
A;Experimental source: ovarian granulosa cells
C;Comment: This receptor belongs to the family of the G-protein coupled receptors. It
ermatogenesis in male and oogenesis in female.
C;Genetics:
A;Gene: fshr
A;Map position: 3 q2.2-q2.3
C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repea
C;Keywords: G protein-coupled receptor; hormone receptor; transmembrane protein
F;1-365/Domain: follicle-stimulating hormone binding #status predicted <HOB>
F;70-94/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>
F;366-388/Domain: transmembrane #status predicted <TM1>
F;398-420/Domain: transmembrane #status predicted <TM2>
F;443-464/Domain: transmembrane #status predicted <TM3>
F;485-507/Domain: transmembrane #status predicted <TM4>
F;528-549/Domain: transmembrane #status predicted <TM5>
F;573-596/Domain: transmembrane #status predicted <TM6>
F;608-629/Domain: transmembrane #status predicted <TM7>

Query Match 24.0%; Score 132; DB 2; Length 694;
Best Local Similarity 26.4%; Pred. No. 1.1e-06;
Matches 28; Conservative 29; Mismatches 41; Indels 8; Gaps 3;

Qy 2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
Db 536 LNVLAFVVICGCYIHIYLTVRNPNIMSS------SSDTKIAKRMAMLIFTDFLCMVPISF 589

Qy 62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTL-TTNFFKD 105
Db 590 FAISASLKVPLITVSKLKILLVLFYPINSCANPFLYAIFTKNFRRD 635

RESULT 14
JC2237
follitropin receptor, testis - horse
N;Alternate names: eFSHR
C;Species: Equus caballus (domestic horse)
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 13-Aug-1999
C;Accession: JC2237; JC2370
R;Robert, P.; Amsellem, S.; Christophe, S.; Benifla, J.L.; Bellet, D.; Koman, A.; Bid
Biochem. Biophys. Res. Commun. 201, 201-207, 1994
A;Title: Cloning and sequencing of the equine testicular follitropin receptor.
A;Reference number: JC2237; MUID:94256980; PMID:8198575
A;Accession: JC2237
A;Molecule type: mRNA
A;Residues: 1-694 <ROB>
A;Cross-references: GB:S70150; NID:g546896; PIDN:AAB30854.1; PID:g546897
A;Experimental source: testis
C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repea
C;Keywords: glycoprotein; hormone receptor; transmembrane protein
F;56-70/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F;194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F;366-386/Domain: transmembrane #status predicted <TM1>
F;398-420/Domain: transmembrane #status predicted <TM2>
F;443-464/Domain: transmembrane #status predicted <TM3>
F;485-507/Domain: transmembrane #status predicted <TM4>
F;528-549/Domain: transmembrane #status predicted <TM5>
F;573-596/Domain: transmembrane #status predicted <TM6>
F;608-629/Domain: transmembrane #status predicted <TM7>
F;191,199,268,293/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 23.5%; Score 129; DB 2; Length 694;
Best Local Similarity 27.4%; Pred. No. 2.3e-06;
Matches 29; Conservative 27; Mismatches 42; Indels 8; Gaps 3;

```
Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|:|||::|   || :: :::   : ::       :  :| |   ::|:| :|  |:
Db     536 LNVLAFVVICGCYIHIYLTVRNPNIVSS------SSDTKIAKRMAILIFTDFLCMAPISF 589

Qy      62 VKILSLFRVEIPDTMTSWI-VIFFLPVNSALNPILYTL-TTNFFKD 105
             | :  :| :     | | :: | |:||  || || : | || :|
Db     590 FAISASLKVPLITVSKSKILLVLFYPINSCANPFLYAIFTKNFRRD 635


RESULT 15
JC7361
follitropin receptor precursor - newt
N;Alternate names: follicle-stimulating hormone receptor
C;Species: Cynops pyrrhogaster (newt)
C;Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 31-Dec-2000
C;Accession: JC7361
R;Nakayama, Y.; Yamamoto, T.; Oba, Y.; Nagahama, Y.; Abe, S.
Biochem. Biophys. Res. Commun. 275, 121-128, 2000
A;Title: Molecular cloning, functional characterization, and gene exppression of a folli
A;Reference number: JC7361
A;Contents: Testis
A;Accession: JC7361
A;Molecule type: mRNA
A;Residues: 1-696 <NAK>
A;Cross-references: DDBJ:AB005587
C;Comment: This protein, containing seven transmembrane domains and a large glycosylated
in and thyroid stimulating hormone. This receptor has a common signal transduction pathw
C;Genetics:
A;Gene: fsh-r
C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
C;Keywords: disulfide bond; glycolysis; glycoprotein; hormone receptor; testis; transmem
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-696/Product: follicle-stimulating hormone receptor #status predicted <MAT>
F;18-359/Domain: extracellular #status predicted <EXT>
F;370-389/Domain: transmembrane #status predicted <TM1>
F;402-424/Domain: transmembrane #status predicted <TM2>
F;447-468/Domain: transmembrane #status predicted <TM3>
F;489-511/Domain: transmembrane #status predicted <TM4>
F;532-553/Domain: transmembrane #status predicted <TM5>
F;577-600/Domain: transmembrane #status predicted <TM6>
F;612-633/Domain: transmembrane #status predicted <TM7>
F;46,190,198,267,292/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;445-520/Disulfide bonds: #status predicted

  Query Match             22.9%;  Score 125.5;  DB 2;  Length 696;
  Best Local Similarity   26.0%;  Pred. No. 5.5e-06;
  Matches   27;  Conservative   26;  Mismatches   44;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|: |||||   || :: :::   : ::       :  :| |   ::|:| :|  |:
Db     540 LNVCAFLIICACYIGIYLTVRNPNVVSS------NSDTKIAKRMAILIFTDFLCMAPISF 593

Qy      62 VKILSLFRVEIPDTMTSWI-VIFFLPVNSALNPILYTLTTNFFK 104
             | :  :: :     | | :: | |:||  || || : |  |:
Db     594 FAISASLKIPLITVSKSKILLVLFYPINSCANPFLYAIFTKTFR 637


Search completed: May 22, 2003, 12:47:53
Job time : 44 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 22, 2003, 12:34:09 ; Search time 25 Seconds
(without alignments)
177.519 Million cell updates/sec

Title: US-09-930-312-2
Perfect score: 549
Sequence: 1 GVNLLAFLIIVFSYITMFCS.....NSALNPILYTLTTNFFKDKL 107

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	549	100.0	754	1	LGR8_HUMAN	Q8wxd0 homo sapien
2	478	87.1	737	1	LGR8_MOUSE	Q91zz5 mus musculu
3	376	68.5	757	1	LGR7_HUMAN	Q9hbx9 homo sapien
4	210	38.3	1115	1	GPCR_LYMST	P46023 lymnaea sta
5	146.5	26.7	319	1	AA3R_MOUSE	Q61618 mus musculu
6	145.5	26.5	320	1	AA3R_RAT	P28647 rattus norv
7	145.5	26.5	701	1	LSHR_BOVIN	Q28005 bos taurus
8	139.5	25.4	696	1	LSHR_PIG	P16582 sus scrofa
9	139	25.3	925	1	GLHR_ANTEL	P35409 anthopleura
10	138.5	25.2	366	1	LSHR_CHICK	Q90674 gallus gall
11	138.5	25.2	676	1	LSHR_CALJA	O02721 callithrix
12	138.5	25.2	699	1	LSHR_HUMAN	P22888 homo sapien
13	136.5	24.9	700	1	LSHR_RAT	P16235 rattus norv
14	132.5	24.1	693	1	FSHR_CHICK	P79763 gallus gall
15	132.5	24.1	700	1	LSHR_MOUSE	P30730 mus musculu
16	129	23.5	694	1	FSHR_HORSE	P47799 equus cabal
17	126	23.0	692	1	FSHR_MOUSE	P35378 mus musculu
18	126	23.0	695	1	FSHR_PIG	P49059 sus scrofa
19	125	22.8	695	1	FSHR_BOVIN	P35376 bos taurus
20	125	22.8	695	1	FSHR_HUMAN	P23945 homo sapien
21	125	22.8	695	1	FSHR_SHEEP	P35379 ovis aries
22	124	22.6	326	1	AA1R_CAVPO	P47745 cavia porce
23	124	22.6	907	1	LGR5_MOUSE	Q9z1p4 mus musculu
24	123	22.4	326	1	AA1R_HUMAN	P30542 homo sapien
25	122	22.2	687	1	FSHR_EQUAS	Q95179 equus asinu
26	122	22.2	692	1	FSHR_RAT	P20395 rattus norv
27	122	22.2	695	1	FSHR_MACFA	P32212 macaca fasc
28	121	22.0	324	1	AA1R_CHICK	P49892 gallus gall
29	120.5	21.9	317	1	AA3R_SHEEP	P35342 ovis aries
30	120.5	21.9	763	1	TSHR_BOVIN	Q27987 bos taurus
31	120.5	21.9	764	1	TSHR_CANFA	P14763 canis famil
32	120.5	21.9	764	1	TSHR_SHEEP	P56495 ovis aries
33	120	21.9	326	1	AA1R_MOUSE	Q60612 mus musculu
34	120	21.9	326	1	AA1R_RAT	P25099 rattus norv
35	120	21.9	828	1	LGR6_HUMAN	Q9hbx8 homo sapien
36	119.5	21.8	318	1	AA3R_HUMAN	P33765 homo sapien
37	119.5	21.8	764	1	TSHR_RAT	P21463 rattus norv
38	119	21.7	326	1	AA1R_CANFA	P11616 canis famil
39	119	21.7	332	1	AA2B_MOUSE	Q60614 mus musculu
40	118.5	21.6	314	1	AA3R_CANFA	Q28309 canis famil
41	118.5	21.6	764	1	TSHR_MOUSE	P47750 mus musculu
42	117.5	21.4	319	1	AA3R_RABIT	O02667 oryctolagus
43	117	21.3	326	1	AA1R_BOVIN	P28190 bos taurus
44	116	21.1	332	1	AA2B_RAT	P29276 rattus norv
45	115	20.9	332	1	AA2B_HUMAN	P29275 homo sapien

ALIGNMENTS

RESULT 1
LGR8_HUMAN
ID LGR8_HUMAN STANDARD; PRT; 754 AA.
AC Q8WXD0;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Relaxin receptor 2 (Leucine-rich repeat-containing G protein-coupled
DE receptor 8) (G protein-coupled receptor affecting testicular descent).
GN LGR8 OR GREAT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND MUTAGENESIS OF ASP-647.
RX MEDLINE=21669315; PubMed=11809971;
RA Hsu S.Y., Nakabayashi K., Nishi S., Kumagai J., Kudo M.,
RA Sherwood O.D., Hsueh A.J.W.;
RT "Activation of orphan receptors by the hormone relaxin.";
RL Science 295:671-674(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Gorlov I.P., Kamat A., Jones E., Lamb D., Truong A., Bogatcheva N.,
RA Bishop C.E., McElreavey K., Agoulnik A.I.;
RT "Mutations of the GREAT gene cause cryptorchidism.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: Receptor for relaxin. The activity of this receptor is
CC mediated by G proteins leading to stimulation of adenylate cyclase
CC and an increase of cAMP. May also be a receptor for Leydig
CC insulin-like peptide (INSL3).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Expressed mainly in the brain kidney, muscle,
CC testis, thyroid,uterus, peripheral blood cells and bone marrow.
CC -!- DISEASE: Defects in LGR8 seems to be a cause of impaired
CC testicular descent (known as cryptorchidism). It is one of the
CC most frequent congenital abnormalities in humans, involving 2-5%
CC of male births. Cryptorchidism is associated with increased risk
CC of infertility and testicular cancer.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -!- SIMILARITY: CONTAINS 1 LDL-RECEPTOR CLASS A DOMAIN.
CC -!- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).
CC -!- CAUTION: It is uncertain whether Met-1 or Met-18 is the initiator.
CC ----------

CC ----------
DR EMBL; AF403384; AAL69324.2; -.
DR EMBL; AF453828; AAL73946.1; -.
DR MIM; 606655; -.
DR MIM; 219050; -.

```
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR002172; LDL_recept_A.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00057; ldl_recept_a; 1.
DR   Pfam; PF00560; LRR; 8.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR00019; LEURICHRPT.
DR   SMART; SM00192; LDLa; 1.
DR   SMART; SM00369; LRR_TYP; 10.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS01209; LDLRA_1; 1.
DR   PROSITE; PS50068; LDLRA_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Repeat;
KW   Leucine-rich repeat.
FT   DOMAIN        1    416       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    417    437       1 (POTENTIAL).
FT   DOMAIN      438    455       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    456    476       2 (POTENTIAL).
FT   DOMAIN      477    507       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    508    528       3 (POTENTIAL).
FT   DOMAIN      529    537       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    538    558       4 (POTENTIAL).
FT   DOMAIN      559    592       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    593    613       5 (POTENTIAL).
FT   DOMAIN      614    639       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    640    660       6 (POTENTIAL).
FT   DOMAIN      661    673       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    674    694       7 (POTENTIAL).
FT   DOMAIN      695    754       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       44     81       LDL-RECEPTOR CLASS A.
FT   REPEAT      136    159       LRR 1.
FT   REPEAT      161    183       LRR 2.
FT   REPEAT      184    207       LRR 3.
FT   REPEAT      209    231       LRR 4.
FT   REPEAT      233    255       LRR 5.
FT   REPEAT      256    279       LRR 6.
FT   REPEAT      281    303       LRR 7.
FT   REPEAT      304    327       LRR 8.
FT   REPEAT      329    351       LRR 9.
FT   REPEAT      352    375       LRR 10.
FT   DISULFID    495    573       BY SIMILARITY.
FT   CARBOHYD     54     54       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    138    138       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    274    274       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    335    335       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    378    378       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   MUTAGEN     647    647       D->Y: LEADS TO CONSTITUTIVE INCREASE OF
FT                                BASAL CAMP.
SQ   SEQUENCE   754 AA;  86452 MW;  2088ECD204C6A6C5 CRC64;

  Query Match           100.0%;  Score 549;  DB 1;  Length 754;
  Best Local Similarity  100.0%;  Pred. No. 3.4e-45;
  Matches  107;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     596 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 655

Qy      61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
           |||||||||||||||||||||||||||||||||||||||||||||||
Db     656 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 702


RESULT 2
LGR8_MOUSE
ID   LGR8_MOUSE     STANDARD;      PRT;   737 AA.
AC   Q91ZZ5;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Relaxin receptor 2 (Leucine-rich repeat-containing G protein-coupled
DE   receptor 8) (G protein-coupled receptor affecting testicular descent)
DE   (G protein-coupled receptor 106).
GN   LGR8 OR GREAT OR GPR106.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6;
RX   MEDLINE=21250990; PubMed=11353515;
RA   Overbeek P.A., Gorlov I.P., Sutherland R.W., Houston J.B.,
RA   Harrison W.R., Boettger-Tong H.L., Bishop C.E., Agoulnik A.I.;
RT   "A transgenic insertion causing cryptorchidism in mice.";
RL   Genesis 30:26-35(2001).
CC   -!- FUNCTION: Receptor for relaxin. The activity of this receptor is
CC       mediated by G proteins leading to stimulation of adenylate cyclase
CC       and an increase of cAMP. May also be a receptor for Leydig
CC       insulin-like peptide (INSL3) (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed in embryonic and adult gonads of
CC       males and females, as well in male gubernarculum. Expressed also
CC       in brain. Not detected in kidney, spleen and heart.
CC   -!- DISEASE: Defects in LGR8 seems to be a cause of impaired
CC       testicular descent (known as cryptorchidism).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   -!- SIMILARITY: CONTAINS 1 LDL-RECEPTOR CLASS A DOMAIN.
CC   -!- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).
CC   -!- MISCELLANEOUS: By homology with the human sequence, it is
CC       uncertain whether Met-1 is the initiator.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF346501; AAL08943.1; -.
DR   MGD; MGI:2153463; Gpr106.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR002172; LDL_recept_A.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00057; ldl_recept_a; 1.
DR   Pfam; PF00560; LRR; 8.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS01209; LDLRA_1; 1.
DR   PROSITE; PS50068; LDLRA_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Repeat;
KW   Leucine-rich repeat.
FT   DOMAIN        1    399       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    400    420       1 (POTENTIAL).
FT   DOMAIN      421    438       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    439    459       2 (POTENTIAL).
FT   DOMAIN      460    490       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    491    511       3 (POTENTIAL).
FT   DOMAIN      512    520       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    521    541       4 (POTENTIAL).
FT   DOMAIN      542    575       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    576    596       5 (POTENTIAL).
FT   DOMAIN      597    622       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    623    643       6 (POTENTIAL).
FT   DOMAIN      644    644       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    645    665       7 (POTENTIAL).
FT   DOMAIN      666    737       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       27     64       LDL-RECEPTOR CLASS A.
FT   REPEAT      119    142       LRR 1.
FT   REPEAT      144    166       LRR 2.
```

```
FT   REPEAT      167    190       LRR 3.
FT   REPEAT      192    214       LRR 4.
FT   REPEAT      215    238       LRR 5.
FT   REPEAT      239    262       LRR 6.
FT   REPEAT      263    286       LRR 7.
FT   REPEAT      287    310       LRR 8.
FT   REPEAT      312    334       LRR 9.
FT   REPEAT      335    358       LRR 10.
FT   DISULFID    478    556       BY SIMILARITY.
FT   CARBOHYD     37     37       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    121    121       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    257    257       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    318    318       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    350    350       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    361    361       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   737 AA;  82943 MW;  AF60F635EA1ACE49 CRC64;

  Query Match             87.1%;  Score 478;  DB 1;  Length 737;
  Best Local Similarity   86.0%;  Pred. No. 1.9e-38;
  Matches   92;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
           |||||||||:|| ||:|||||| |||||| |||:  |:||||||||||||||||||||||
Db     579 GVNLLAFLVIVISYVTMFCSIHKTALQTAEVRSHIGKEVAVANRFFFIVFSDAICWIPVF 638

Qy      61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
           ||||||| :|||| |:|||||:|||||||||||||||||:||||||
Db     639 VVKILSLLQVEIPGTITSWIVVFFLPVNSALNPILYTLTTSFFKDKL 685


RESULT 3
LGR7_HUMAN
ID   LGR7_HUMAN     STANDARD;      PRT;   757 AA.
AC   Q9HBX9;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Relaxin receptor 1 (Leucine-rich repeat-containing G protein-coupled
DE   receptor 7).
GN   LGR7.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND MUTAGENESIS OF ASP-637.
RX   MEDLINE=20388592; PubMed=10935549;
RA   Hsu S.Y., Kudo M., Chen T., Nakabayashi K., Bhalla A.,
RA   van der Spek P.J., van Duin M., Hsueh A.J.W.;
RT   "The three subfamilies of leucine-rich repeat-containing G protein-
RT   coupled receptors (LGR): identification of LGR6 and LGR7 and the
RT   signaling mechanism for LGR7.";
RL   Mol. Endocrinol. 14:1257-1271(2000).
RN   [2]
RP   CHARACTERIZATION.
RX   PubMed=11517286;
RA   Bartsch O., Bartlick B., Ivell R.;
RT   "Relaxin signalling links tyrosine phosphorylation to
RT   phosphodiesterase and adenylyl cyclase activity.";
RL   Mol. Hum. Reprod. 7:799-809(2001).
CC   -!- FUNCTION: Receptor for relaxin. The activity of this receptor is
CC       mediated by G proteins leading to stimulation of adenylate cyclase
CC       and an increase of cAMP. Binding of the ligand may also
CC       activate a tyrosine kinase pathway that inhibits the activity of a
CC       phosphodiesterase that degrades cAMP.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
CC       produced by alternative splicing.
CC   -!- TISSUE SPECIFICITY: Expressed in the brain, kidney, testis,
CC       placenta, uterus, ovary, adrenal, prostate, skin and heart. Not
CC       detected in spleen.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   -!- SIMILARITY: CONTAINS 1 LDL-RECEPTOR CLASS A DOMAIN.
CC   -!- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF190500; AAG17167.1; -.
DR   MIM; 606654; -.
DR   HSSP; P01130; 1LDR.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR002172; LDL_recept_A.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00057; ldl_recept_a; 1.
DR   Pfam; PF00560; LRR; 8.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   SMART; SM00192; LDLa; 1.
DR   SMART; SM00370; LRR; 7.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 10.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS01209; LDLRA_1; 1.
DR   PROSITE; PS50068; LDLRA_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Repeat;
KW   Leucine-rich repeat; Alternative splicing.
FT   DOMAIN        1    409       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    410    430       1 (POTENTIAL).
FT   DOMAIN      431    443       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    444    464       2 (POTENTIAL).
FT   DOMAIN      465    486       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    487    507       3 (POTENTIAL).
FT   DOMAIN      508    527       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    528    548       4 (POTENTIAL).
FT   DOMAIN      549    577       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    578    598       5 (POTENTIAL).
FT   DOMAIN      599    629       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    630    650       6 (POTENTIAL).
FT   DOMAIN      651    651       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    652    672       7 (POTENTIAL).
FT   DOMAIN      673    757       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       26     63       LDL-RECEPTOR CLASS A.
FT   REPEAT      125    148       LRR 1.
FT   REPEAT      150    172       LRR 2.
FT   REPEAT      173    196       LRR 3.
FT   REPEAT      198    220       LRR 4.
FT   REPEAT      221    244       LRR 5.
FT   REPEAT      246    269       LRR 6.
FT   REPEAT      270    293       LRR 7.
FT   REPEAT      294    317       LRR 8.
FT   REPEAT      319    341       LRR 9.
FT   REPEAT      343    365       LRR 10.
FT   DISULFID    485    563       BY SIMILARITY.
FT   CARBOHYD     36     36       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    127    127       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    264    264       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    272    272       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    325    325       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    368    368       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC     63     96       GDNNGWSMQFDKYFASYYKMTSQYPFEAETPECL -> V
FT                                 (IN ISOFORM 2).
FT   MUTAGEN     637    637       D->Y: LEADS TO CONSTITUTIVE INCREASE OF
FT                                 BASAL CAMP.
SQ   SEQUENCE   757 AA;  86992 MW;  8079E8DBF3A3EF21 CRC64;

  Query Match             68.5%;  Score 376;  DB 1;  Length 757;
```

```
  Best Local Similarity   64.5%;  Pred. No. 1e-28;
  Matches   69;  Conservative   20;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 GVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
           |:|| ||:|||||| :|| |: ::|:  ||:||   :|: :| |||||||:||:||||:|
Db     586 GINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVFTDALCWIPIF 645

Qy      61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
           ||| ||| :|||| |:|||:||| ||:|||||||||||||  ||: :
Db     646 VVKFLSLLQVEIPGTITSWVVIFILPINSALNPILYTLTTRPFKEMI 692


RESULT 4
GPCR_LYMST
ID   GPCR_LYMST     STANDARD;      PRT;  1115 AA.
AC   P46023;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   G-protein coupled receptor GRL101 precursor.
OS   Lymnaea stagnalis (Great pond snail).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
OC   Lymnaeidae; Lymnaea.
OX   NCBI_TaxID=6523;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=CNS;
RX   MEDLINE=94255418; PubMed=8197140;
RA   Tensen C.P., van Kesteren E.R., Planta R.J., Cox K.J.A., Burke J.F.,
RA   van Heerikhuizen H., Vreugdenhil E.;
RT   "A G protein-coupled receptor with low density lipoprotein-binding
RT   motifs suggests a role for lipoproteins in G-linked signal
RT   transduction.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:4816-4820(1994).
CC   -!- FUNCTION: MIGHT DIRECTLY TRANSDUCE SIGNALS CARRIED BY LARGE
CC       EXTRACELLULAR (LIPO)PROTEIN(COMPLEXE)S INTO NEURONAL EVENTS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN A SMALL NUMBER OF
CC       NEURONS WITHIN THE CENTRAL NERVOUS SYSTEM AND TO A LESSER EXTENT
CC       IN THE HEART.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   -!- SIMILARITY: CONTAINS 12 LDL-RECEPTOR CLASS A DOMAINS.
CC   -!- SIMILARITY: CONTAINS 7 LEUCINE-RICH REPEATS (LRR).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z23104; CAA80651.1; -.
DR   PIR; S40241; S40241.
DR   HSSP; Q07954; 1CR8.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR002172; LDL_recept_A.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00057; ldl_recept_a; 11.
DR   Pfam; PF00560; LRR; 6.
DR   Pfam; PF01462; LRRNT; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR00261; LDLRECEPTOR.
DR   SMART; SM00192; LDLa; 12.
DR   SMART; SM00370; LRR; 2.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS01209; LDLRA_1; 6.
DR   PROSITE; PS50068; LDLRA_2; 11.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Repeat;
KW   Leucine-rich repeat; Signal.
FT   SIGNAL       1     24       POTENTIAL.
FT   CHAIN       25   1115       G-PROTEIN COUPLED RECEPTOR GRL101.
FT   DOMAIN      25    767       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   768    788       1 (POTENTIAL).
FT   DOMAIN     789    801       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   802    822       2 (POTENTIAL).
FT   DOMAIN     823    857       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   858    878       3 (POTENTIAL).
FT   DOMAIN     879    887       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   888    908       4 (POTENTIAL).
FT   DOMAIN     909    941       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   942    962       5 (POTENTIAL).
FT   DOMAIN     963    988       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   989   1009       6 (POTENTIAL).
FT   DOMAIN    1010   1017       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM  1018   1038       7 (POTENTIAL).
FT   DOMAIN    1039   1115       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      32    523       12 X 40 AA APPROXIMATE TANDEM REPEATS
FT                               SIMILAR TO THE LDL-RECEPTOR CLASS A.
FT   DOMAIN      36     79       LDL-RECEPTOR CLASS A 1.
FT   DOMAIN      77    115       LDL-RECEPTOR CLASS A 2.
FT   DOMAIN     116    155       LDL-RECEPTOR CLASS A 3.
FT   DOMAIN     156    196       LDL-RECEPTOR CLASS A 4.
FT   DOMAIN     195    232       LDL-RECEPTOR CLASS A 5.
FT   DOMAIN     231    269       LDL-RECEPTOR CLASS A 6.
FT   DOMAIN     272    318       LDL-RECEPTOR CLASS A 7.
FT   DOMAIN     320    363       LDL-RECEPTOR CLASS A 8.
FT   DOMAIN     365    403       LDL-RECEPTOR CLASS A 9.
FT   DOMAIN     404    442       LDL-RECEPTOR CLASS A 10.
FT   DOMAIN     444    485       LDL-RECEPTOR CLASS A 11.
FT   DOMAIN     486    525       LDL-RECEPTOR CLASS A 12.
FT   REPEAT     557    579       LRR 1.
FT   REPEAT     582    605       LRR 2.
FT   REPEAT     607    629       LRR 3.
FT   REPEAT     630    653       LRR 4.
FT   REPEAT     655    677       LRR 5.
FT   REPEAT     678    701       LRR 6.
FT   REPEAT     703    725       LRR 7.
FT   DISULFID    38     53       BY SIMILARITY.
FT   DISULFID    46     66       BY SIMILARITY.
FT   DISULFID    60     77       BY SIMILARITY.
FT   DISULFID    79     91       BY SIMILARITY.
FT   DISULFID    86    104       BY SIMILARITY.
FT   DISULFID    98    113       BY SIMILARITY.
FT   DISULFID   118    131       BY SIMILARITY.
FT   DISULFID   138    153       BY SIMILARITY.
FT   DISULFID   158    170       BY SIMILARITY.
FT   DISULFID   165    183       BY SIMILARITY.
FT   DISULFID   177    194       BY SIMILARITY.
FT   DISULFID   202    220       BY SIMILARITY.
FT   DISULFID   214    230       BY SIMILARITY.
FT   DISULFID   233    245       BY SIMILARITY.
FT   DISULFID   240    258       BY SIMILARITY.
FT   DISULFID   252    267       BY SIMILARITY.
FT   DISULFID   274    291       BY SIMILARITY.
FT   DISULFID   282    304       BY SIMILARITY.
FT   DISULFID   298    316       BY SIMILARITY.
FT   DISULFID   322    339       BY SIMILARITY.
FT   DISULFID   334    352       BY SIMILARITY.
FT   DISULFID   346    361       BY SIMILARITY.
FT   DISULFID   367    379       BY SIMILARITY.
FT   DISULFID   374    392       BY SIMILARITY.
FT   DISULFID   386    401       BY SIMILARITY.
FT   DISULFID   406    418       BY SIMILARITY.
FT   DISULFID   413    431       BY SIMILARITY.
FT   DISULFID   425    440       BY SIMILARITY.
FT   DISULFID   446    458       BY SIMILARITY.
FT   DISULFID   453    474       BY SIMILARITY.
FT   DISULFID   465    483       BY SIMILARITY.
```

```
FT   DISULFID    488    500       BY SIMILARITY.
FT   DISULFID    495    513       BY SIMILARITY.
FT   DISULFID    507    523       BY SIMILARITY.
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    166    166       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    269    269       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    318    318       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    482    482       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    502    502       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    571    571       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    618    618       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    624    624       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    685    685       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1115 AA;  125865 MW;  E9BB01297ECE356C CRC64;

  Query Match           38.3%;  Score 210;  DB 1;  Length 1115;
  Best Local Similarity   41.6%;  Pred. No. 9.5e-13;
  Matches   42;  Conservative   23;  Mismatches   30;  Indels    6;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKT--ALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPV 59
           :|||:|::|  ||: ||   :||  |::| | :|    : |:| |   || :|  ||:|:
Db     946 LNLLSFVLIASSYLWMFSVAKKTRSAVRTAESKN----DNAMARRMTLIVMTDFCCWVPI 1001

Qy      60 FVVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTT 100
            |:  :||      | : :|| :| ||:||| ||::|||:|
Db    1002 IVLGFVSLAGARADDQVYAWIAVFVLPLNSATNPVIYTLST 1042


RESULT 5
AA3R_MOUSE
ID   AA3R_MOUSE     STANDARD;      PRT;   319 AA.
AC   Q61618; Q9R202;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Adenosine 3 receptor (A3AR).
GN   ADORA3 OR GPCR2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99208672; PubMed=10191095;
RA   Zhao Z., Francis C., Ravid K.;
RT   "Characterization of the mouse A3 adenosine receptor gene: exon/intron
RT   organization and promoter activity.";
RL   Genomics 57:152-155(1999).
RN   [2]
RP   SEQUENCE OF 111-239 FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=94116980; PubMed=8288218;
RA   Wilkie T.M., Chen Y., Gilbert D.J., Moore K.J., Yu L., Simon M.I.,
RA   Copeland N.G., Jenkins N.A.;
RT   "Identification, chromosomal location, and genome organization of
RT   mammalian G-protein-coupled receptors.";
RL   Genomics 18:175-184(1993).
CC   -!- FUNCTION: RECEPTOR FOR ADENOSINE. THE ACTIVITY OF THIS RECEPTOR
CC       IS MEDIATED BY G PROTEINS WHICH INHIBITS ADENYLYL CYCLASE.
CC       POSSIBLE ROLE IN REPRODUCTION.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF069778; AAC82643.1; -.
DR   EMBL; L20331; AAA16851.1; -.
DR   MGD; MGI:104847; Adora3.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Lipoprotein; Palmitate.
FT   DOMAIN        1     15       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     16     38       1 (POTENTIAL).
FT   DOMAIN       39     49       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     50     73       2 (POTENTIAL).
FT   DOMAIN       74     85       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     86    107       3 (POTENTIAL).
FT   DOMAIN      108    127       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    128    149       4 (POTENTIAL).
FT   DOMAIN      150    178       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    179    199       5 (POTENTIAL).
FT   DOMAIN      200    232       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    233    256       6 (POTENTIAL).
FT   DOMAIN      257    262       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    263    285       7 (POTENTIAL).
FT   DOMAIN      286    319       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      5      5       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     13     13       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    161    161       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     84    167       BY SIMILARITY.
FT   LIPID       304    304       PALMITATE (POTENTIAL).
FT   CONFLICT    184    184       I -> V (IN REF. 2).
FT   CONFLICT    196    196       I -> V (IN REF. 2).
FT   CONFLICT    214    214       T -> S (IN REF. 2).
SQ   SEQUENCE   319 AA;  36448 MW;  6AD7F46A9F03D9CC CRC64;

  Query Match           26.7%;  Score 146.5;  DB 1;  Length 319;
  Best Local Similarity   30.2%;  Pred. No. 3.5e-07;
  Matches   32;  Conservative   22;  Mismatches   47;  Indels    5;  Gaps    2;

Qy       4 LLAFLIIVFSYITMFCSIQKTALQT----TEVRNCFGREVAVANRFFFIVFSDAICWIPV 59
           |:  :::   |: :|  |:    |      | |  :|||   |   | ::|  |:||:|:
Db     188 LVPLVVMCIIYLDIFYIIRNKLSQNLTGFRETRAFYGREFKTAKSLFLVLFLFALCWLPL 247

Qy      60 FVVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKD 105
            ::  :| | |:||| :   : |     || :|||:|      ||:
Db     248 SIINFVSYFDVKIPD-VAMCLGILLSHANSMMNPIVYACKIKKFKE 292


RESULT 6
AA3R_RAT
ID   AA3R_RAT       STANDARD;      PRT;   320 AA.
AC   P28647; Q63792;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Adenosine A3 receptor (TGPCR1).
GN   ADORA3.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC   TISSUE=Brain;
RX   MEDLINE=92366475; PubMed=1323836;
RA   Zhou Q.Y., Li C.Y., Olah M.E., Johnson R.A., Stiles G.L., Civelli O.;
RT   "Molecular cloning and characterization of an adenosine receptor: the
RT   A3 adenosine receptor.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:7432-7436(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Testis;
RX   MEDLINE=91285122; PubMed=1647979;
RA   Meyerhof W., Mueller-Brechlin R., Richter D.;
RT   "Molecular cloning of a novel putative G-protein coupled receptor
RT   expressed during rat spermiogenesis.";
```

```
RL   FEBS Lett. 284:155-160(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=96196578; PubMed=8612733;
RA   Sajjadi F.G., Boyle D.L., Domingo R.C., Firestein G.S.;
RT   "cDNA cloning and characterization of A3i, an alternatively spliced
RT   rat A3 adenosine receptor variant.";
RL   FEBS Lett. 382:125-129(1996).
CC   -!- FUNCTION: RECEPTOR FOR ADENOSINE. THE ACTIVITY OF THIS RECEPTOR
CC       IS MEDIATED BY G PROTEINS WHICH INHIBITS ADENYLYL CYCLASE.
CC       POSSIBLE ROLE IN REPRODUCTION.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS: TWO ISOFORMS ARE PRODUCED BY ALTERNATIVE
CC       SPLICING.
CC   -!- TISSUE SPECIFICITY: TESTIS, PARTICULARLY IN SPERMATOCYTES AND
CC       SPERMATIDS BUT NOT IN SPERMATOGONIA. LOW LEVELS IN THE BRAIN.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED DURING SPERMIOGENESIS.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X59249; CAA41937.1; -.
DR   EMBL; M94152; AAA40680.1; -.
DR   EMBL; X93219; CAA63702.1; -.
DR   PIR; S17177; S17177.
DR   PIR; A46152; A46152.
DR   HSSP; P29274; 1MMH.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Lipoprotein; Palmitate; Alternative splicing.
FT   DOMAIN        1     16       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     17     39       1 (POTENTIAL).
FT   DOMAIN       40     50       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     51     74       2 (POTENTIAL).
FT   DOMAIN       75     86       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     87    108       3 (POTENTIAL).
FT   DOMAIN      109    128       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    129    150       4 (POTENTIAL).
FT   DOMAIN      151    179       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    180    200       5 (POTENTIAL).
FT   DOMAIN      201    233       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    234    257       6 (POTENTIAL).
FT   DOMAIN      258    263       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    264    286       7 (POTENTIAL).
FT   DOMAIN      287    320       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      4      4       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      5      5       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     85    168       BY SIMILARITY.
FT   LIPID       305    305       PALMITATE (POTENTIAL).
FT   VARSPLIC    119    119       R -> RLSFLVVNLPFSSPHSNR (IN ISOFORM A3I).
FT   CONFLICT     18     18       I -> V (IN REF. 1).
FT   CONFLICT     68     68       I -> V (IN REF. 3).
FT   CONFLICT     74     83       VSLEVQMHFY -> SAWRSRCTSM (IN REF. 2).
FT   CONFLICT    132    132       F -> L (IN REF. 3).
FT   CONFLICT    172    172       S -> F (IN REF. 1).
FT   CONFLICT    288    288       I -> N (IN REF. 1).
FT   CONFLICT    291    295       FKETY -> VQRNH (IN REF. 1).
SQ   SEQUENCE   320 AA;  36629 MW;  59DBA40C594ACDAB CRC64;

  Query Match           26.5%;  Score 145.5;  DB 1;  Length 320;
  Best Local Similarity 29.2%;  Pred. No. 4.4e-07;
  Matches   31;  Conservative   23;  Mismatches   47;  Indels    5;  Gaps    2;

Qy       4 LLAFLIIVFSYITMFCSIQKTALQT----TEVRNCFGREVAVANRFFFIVFSDAICWIPV 59
           |:  :::   |: :|  |:    |      | |  :|||   |   | ::|  |:||:|:
Db     189 LIPLVVMCIIYLDIFYIIRNKLSQNLTGFRETRAFYGREFKTAKSLFLVLFLFALCWLPL 248

Qy      60 FVVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKD 105
             ::  :| | |:||: :   : |     || :|||:|      ||:
Db     249 SIINFVSYFNVKIPE-IAMCLGILLSHANSMMNPIVYACKIKKFKE 293


RESULT 7
LSHR_BOVIN
ID   LSHR_BOVIN     STANDARD;      PRT;   701 AA.
AC   Q28005; P79133;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Lutropin-choriogonadotropic hormone receptor precursor (LH/CG-R)
DE   (LSH-R) (Luteinizing hormone receptor).
GN   LHCGR.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Holstein; TISSUE=Ovary, and Testis;
RA   Lussier J.G., Houde A., Ethier J., Silversides D.W.;
RL   Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 80-701 FROM N.A.
RX   MEDLINE=98193601; PubMed=9532424;
RA   Mamluk R., Wolfenson D., Meidan R.;
RT   "LH receptor mRNA and cytochrome P450 side-chain cleavage expression
RT   in bovine theca and granulosa cells luteinized by LH or forskolin.";
RL   Domest. Anim. Endocrinol. 15:103-114(1998).
CC   -!- FUNCTION: RECEPTOR FOR LUTROPIN-CHORIOGONADOTROPIC HORMONE.
CC       THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH
CC       ACTIVATE ADENYLATE CYCLASE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC       SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       FSH/LSH/TSH SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U20504; AAC24012.1; -.
DR   EMBL; U87230; AAC33486.1; -.
DR   HSSP; P22888; 1LUT.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 2.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing.
FT   SIGNAL        1     26       POTENTIAL.
FT   CHAIN        27    701       LUTROPIN-CHORIOGONADOTROPIC HORMONE
FT                                RECEPTOR.
FT   DOMAIN       27    365       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    366    387       1 (POTENTIAL).
```

```
FT   DOMAIN      388    397       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    398    418       2 (POTENTIAL).
FT   DOMAIN      419    441       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    442    464       3 (POTENTIAL).
FT   DOMAIN      465    484       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    485    507       4 (POTENTIAL).
FT   DOMAIN      508    527       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    528    551       5 (POTENTIAL).
FT   DOMAIN      552    572       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    573    596       6 (POTENTIAL).
FT   DOMAIN      597    607       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    608    629       7 (POTENTIAL).
FT   DOMAIN      630    701       CYTOPLASMIC (POTENTIAL).
FT   REPEAT       50     73       LRR 1.
FT   REPEAT      124    149       LRR 2.
FT   REPEAT      151    173       LRR 3.
FT   REPEAT      174    198       LRR 4.
FT   REPEAT      200    222       LRR 5.
FT   REPEAT      223    246       LRR 6.
FT   DISULFID    441    516       BY SIMILARITY.
FT   CARBOHYD    101    101       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    176    176       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    197    197       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    293    293       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    301    301       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    315    315       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    229    291       MISSING (IN SHORT ISOFORM).
FT   CONFLICT    557    557       N -> D (IN REF. 2).
FT   CONFLICT    577    577       I -> T (IN REF. 2).
FT   CONFLICT    589    589       F -> S (IN REF. 2).
SQ   SEQUENCE   701 AA;  78455 MW;  8CDCE92F59EDBBD7 CRC64;

  Query Match             26.5%;  Score 145.5;  DB 1;  Length 701;
  Best Local Similarity   29.8%;  Pred. No. 8.9e-07;
  Matches   31;  Conservative   24;  Mismatches   42;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|::||:||   || :: ::|   |  |       ::  :| :   ::|:|  |  |:
Db     536 LNVMAFIIICACYIKIYFAVQNPELMAT------NKDTKIAKKMAVLIFTDFTCMAPISF 589

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
             | : |:|  |  | :  ::: | ||||  || || : |  |:
Db     590 FAISAAFKVPLITVTNSKVLLVLFYPVNSCANPFLYAIFTKAFQ 633


RESULT 8
LSHR_PIG
ID   LSHR_PIG       STANDARD;      PRT;   696 AA.
AC   P16582;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Lutropin-choriogonadotropic hormone receptor precursor (LH/CG-R)
DE   (LSH-R) (Luteinizing hormone receptor).
GN   LHCGR.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89332517; PubMed=2502844;
RA   Loosfelt H., Misrahi M., Atger M., Salesse R., Thi M.T.V.H.-L.,
RA   Jolivet A., Guiochon-Mantel A., Sar S., Jallal B., Garnier J.,
RA   Milgrom E.;
RT   "Cloning and sequencing of porcine LH-hCG receptor cDNA: variants
RT   lacking transmembrane domain.";
RL   Science 245:525-528(1989).
CC   -!- FUNCTION: RECEPTOR FOR LUTROPIN-CHORIOGONADOTROPIC HORMONE.
CC       THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH
CC       ACTIVATE ADENYLATE CYCLASE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS: 4 ISOFORMS; A (SHOWN HERE), B, C AND D; ARE
CC       PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       FSH/LSH/TSH SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M29525; AAA31062.1; -.
DR   EMBL; M29526; AAA31063.1; -.
DR   EMBL; M29527; AAA31064.1; -.
DR   EMBL; M29528; AAA31065.1; -.
DR   PIR; A41344; A41344.
DR   PIR; B41344; B41344.
DR   PIR; C41344; C41344.
DR   PIR; D41344; D41344.
DR   HSSP; P22888; 1LUT.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 2.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing.
FT   SIGNAL        1     27       POTENTIAL.
FT   CHAIN        28    696       LUTROPIN-CHORIOGONADOTROPIC HORMONE
FT                                RECEPTOR.
FT   DOMAIN        28    358       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     359    386       1 (POTENTIAL).
FT   DOMAIN       387    395       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     396    418       2 (POTENTIAL).
FT   DOMAIN       419    439       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     440    462       3 (POTENTIAL).
FT   DOMAIN       463    482       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     483    505       4 (POTENTIAL).
FT   DOMAIN       506    525       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     526    547       5 (POTENTIAL).
FT   DOMAIN       548    570       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     571    594       6 (POTENTIAL).
FT   DOMAIN       595    605       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     606    626       7 (POTENTIAL).
FT   DOMAIN       627    696       CYTOPLASMIC (POTENTIAL).
FT   REPEAT        48     71       LRR 1.
FT   REPEAT       122    147       LRR 2.
FT   REPEAT       149    171       LRR 3.
FT   REPEAT       172    196       LRR 4.
FT   REPEAT       198    220       LRR 5.
FT   REPEAT       221    244       LRR 6.
FT   DISULFID     439    514       BY SIMILARITY.
FT   CARBOHYD      99     99       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     174    174       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     195    195       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     291    291       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     299    299       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     313    313       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC     317    329       YSAIFAESELSDW -> LLHGALPATHCLS (IN
FT                                ISOFORM B).
FT   VARSPLIC     330    696       MISSING (IN ISOFORM B).
FT   VARSPLIC     317    331       YSAIFAESELSDWDY -> SKSRADGYQQRHKDC (IN
FT                                ISOFORM C).
FT   VARSPLIC     332    696       MISSING (IN ISOFORM C).
FT   VARSPLIC     317    628       MISSING (IN ISOFORM D).
SQ   SEQUENCE   696 AA;  78092 MW;  593DEF1C25F982FE CRC64;

  Query Match             25.4%;  Score 139.5;  DB 1;  Length 696;
```

```
  Best Local Similarity   28.8%;  Pred. No. 3.3e-06;
  Matches   30;  Conservative   24;  Mismatches   43;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|::||:||   || :: ::|   |  |       ::  :| :   ::|:|  |  |:
Db     534 LNVVAFIIICACYIKIYFAVQNPELMAT------NKDTKIAKKMAVLIFTDFTCMAPISF 587

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
             |  :  :|  |  | :  ::: | ||||  || || : |  |:
Db     588 FAISAALKVPLITVTNSKVLLVLFYPVNSCANPFLYAIFTKAFR 631


RESULT 9
GLHR_ANTEL
ID   GLHR_ANTEL     STANDARD;      PRT;   925 AA.
AC   P35409;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable glycoprotein hormone G-protein coupled receptor precursor.
OS   Anthopleura elegantissima (Sea anemone).
OC   Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC   Nynantheae; Actiniidae; Anthopleura.
OX   NCBI_TaxID=6110;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94107299; PubMed=8280121;
RA   Nothacker H.-P., Grimmelikhuijzen C.J.P.;
RT   "Molecular cloning of a novel, putative G protein-coupled receptor
RT   from sea anemones structurally related to members of the FSH, TSH,
RT   LH/CG receptor family from mammals.";
RL   Biochem. Biophys. Res. Commun. 197:1062-1069(1993).
CC   -!- FUNCTION: PROBABLE RECEPTOR FOR A GLYCOPROTEIN HORMONE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       FSH/LSH/TSH SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z28332; CAA82186.1; -.
DR   PIR; S41908; S41908.
DR   PIR; JC2033; JC2033.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR000832; GPCR_secretin.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR003592; LRR_out.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 6.
DR   SMART; SM00370; LRR; 2.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Alternative splicing; Repeat.
FT   SIGNAL        1     27       POTENTIAL.
FT   CHAIN        28    925       PROBABLE GLYCOPROTEIN HORMONE G-PROTEIN
FT                                COUPLED RECEPTOR.
FT   DOMAIN        28    529       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     530    551       1 (POTENTIAL).
FT   DOMAIN       552    561       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     562    584       2 (POTENTIAL).
FT   DOMAIN       585    606       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     607    628       3 (POTENTIAL).
FT   DOMAIN       629    651       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     652    673       4 (POTENTIAL).
FT   DOMAIN       674    691       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     692    712       5 (POTENTIAL).
FT   DOMAIN       713    739       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     740    763       6 (POTENTIAL).
FT   DOMAIN       764    774       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     775    795       7 (POTENTIAL).
FT   DOMAIN       796    925       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       333    461       5 X APPROXIMATE TANDEM REPEATS.
FT   REPEAT       333    349       1 (INCOMPLETE).
FT   REPEAT       350    384       2.
FT   REPEAT       385    419       3.
FT   REPEAT       420    453       4.
FT   REPEAT       454    461       5 (INCOMPLETE).
FT   CARBOHYD      61     61       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     152    152       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     212    212       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     435    435       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     495    495       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC     235    925       MISSING (IN SHORT ISOFORM).
SQ   SEQUENCE   925 AA;  100059 MW;  D03A256368452FBD CRC64;

  Query Match             25.3%;  Score 139;  DB 1;  Length 925;
  Best Local Similarity   29.6%;  Pred. No. 4.7e-06;
  Matches   32;  Conservative   24;  Mismatches   46;  Indels    6;  Gaps    3;

Qy       2 VNLLAFLIIVFSYITM-FCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
           ||  :|: ::: |  | :  :    ::    ||    :  || |   :||:| :|| |:
Db     700 VNGASFISVMYLYSRMLYVVVSGGDMEGAPKRN----DSKVAKRMAILVFTDMLCWAPIA 755

Qy      61 VVKILSLF-RVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
             :|: | :  :  | :  :::|| |:||  || ||   |  || :|
Db     756 FFGLLAAFGQTLLTVTQSKILLVFFFPINSICNPFLYAFFTKAFKREL 803


RESULT 10
LSHR_CHICK
ID   LSHR_CHICK     STANDARD;      PRT;   366 AA.
AC   Q90674;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Lutropin-choriogonadotropic hormone receptor (LH/CG-R) (LSH-R)
DE   (Luteinizing hormone receptor) (Fragment).
GN   LHCGR.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=96426554; PubMed=8828833;
RA   Johnson A.L., Bridgham J.T., Wagner B.;
RT   "Characterization of a chicken luteinizing hormone receptor (cLH-R)
RT   complementary deoxyribonucleic acid, and expression of cLH-R
RT   messenger ribonucleic acid in the ovary.";
RL   Biol. Reprod. 55:304-309(1996).
CC   -!- FUNCTION: RECEPTOR FOR LUTROPIN-CHORIOGONADOTROPIC HORMONE.
CC       THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH
CC       ACTIVATE ADENYLATE CYCLASE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       FSH/LSH/TSH SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U31987; AAC59907.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
```

```
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Phosphorylation.
FT   NON_TER       1      1
FT   DOMAIN       <1     26       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     27     48       1 (POTENTIAL).
FT   DOMAIN       49     58       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     59     79       2 (POTENTIAL).
FT   DOMAIN       80    102       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    103    125       3 (POTENTIAL).
FT   DOMAIN      126    145       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    146    168       4 (POTENTIAL).
FT   DOMAIN      169    188       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    189    212       5 (POTENTIAL).
FT   DOMAIN      213    233       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    234    257       6 (POTENTIAL).
FT   DOMAIN      258    268       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    269    290       7 (POTENTIAL).
FT   DOMAIN      291    366       CYTOPLASMIC (POTENTIAL).
FT   DISULFID    102    177       BY SIMILARITY.
SQ   SEQUENCE   366 AA;  40759 MW;  C2F7199F9E3CBCCA CRC64;

  Query Match             25.2%;  Score 138.5;  DB 1;  Length 366;
  Best Local Similarity   28.8%;  Pred. No. 2.3e-06;
  Matches   30;  Conservative   23;  Mismatches   44;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|::|||:|   || :: ::|   |          ::  :| |   ::|:|  |  |:
Db     197 LNVIAFLVICACYIKIYVAVQNPELVAA------NKDTKIAKRMAILIFTDFTCMAPISF 250

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
             | :  :|  |  | :  ::: | ||||  || || : |  |:
Db     251 FAISAAIKVPLITVTNSKILLVLFYPVNSCANPFLYAIFTKAFQ 294


RESULT 11
LSHR_CALJA
ID   LSHR_CALJA     STANDARD;      PRT;   676 AA.
AC   O02721;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Lutropin-choriogonadotropic hormone receptor precursor (LH/CG-R)
DE   (LSH-R) (Luteinizing hormone receptor).
GN   LHCGR.
OS   Callithrix jacchus (Common marmoset).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
OC   Callithrix.
OX   NCBI_TaxID=9483;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=97307697; PubMed=9165039;
RA   Zhang F.-P., Rannikko A.S., Manna P.R., Fraser H.M., Huhtaniemi I.T.;
RT   "Cloning and functional expression of the luteinizing hormone
RT   receptor complementary deoxyribonucleic acid from the marmoset monkey
RT   testis: absence of sequences encoding exon 10 in other species.";
RL   Endocrinology 138:2481-2490(1997).
CC   -!- FUNCTION: RECEPTOR FOR LUTROPIN-CHORIOGONADOTROPIC HORMONE.
CC       THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH
CC       ACTIVATE ADENYLATE CYCLASE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       FSH/LSH/TSH SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 5 LEUCINE-RICH REPEATS (LRR).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U80673; AAB53698.1; -.
DR   HSSP; P22888; 1LUT.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 2.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Phosphorylation; Repeat; Leucine-rich repeat.
FT   SIGNAL        1     29       POTENTIAL.
FT   CHAIN        30    676       LUTROPIN-CHORIOGONADOTROPIC HORMONE
FT                                RECEPTOR.
FT   DOMAIN       30    340       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    341    362       1 (POTENTIAL).
FT   DOMAIN      363    372       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    373    393       2 (POTENTIAL).
FT   DOMAIN      394    416       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    417    439       3 (POTENTIAL).
FT   DOMAIN      440    459       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    460    482       4 (POTENTIAL).
FT   DOMAIN      483    502       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    503    526       5 (POTENTIAL).
FT   DOMAIN      527    547       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    548    571       6 (POTENTIAL).
FT   DOMAIN      572    582       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    583    604       7 (POTENTIAL).
FT   DOMAIN      605    676       CYTOPLASMIC (POTENTIAL).
FT   REPEAT       52     75       LRR 1.
FT   REPEAT      126    151       LRR 2.
FT   REPEAT      153    175       LRR 3.
FT   REPEAT      176    200       LRR 4.
FT   REPEAT      225    248       LRR 5.
FT   DISULFID    416    491       BY SIMILARITY.
FT   CARBOHYD    103    103       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    178    178       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    199    199       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   676 AA;  75677 MW;  FB018227641654E1 CRC64;

  Query Match             25.2%;  Score 138.5;  DB 1;  Length 676;
  Best Local Similarity   26.9%;  Pred. No. 4e-06;
  Matches   28;  Conservative   27;  Mismatches   42;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|::||:||   || :: :::   |  |       ::  :| :   ::|:|  |  |:
Db     511 LNVVAFIIICACYIKIYFAVRNPELMAT------NKDTKIAKKMAILIFTDFTCMAPISF 564

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
             | : |::  |  | :  ::: | |:||  || || : |  |:
Db     565 FAISAAFKMPLITVTNSKVLLVLFYPINSCANPFLYAIFTKTFR 608


RESULT 12
LSHR_HUMAN
ID   LSHR_HUMAN     STANDARD;      PRT;   699 AA.
AC   P22888;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Lutropin-choriogonadotropic hormone receptor precursor (LH/CG-R)
DE   (LSH-R) (Luteinizing hormone receptor).
GN   LHCGR OR LHRHR OR LCGR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
```

RC TISSUE=Ovary;
RX MEDLINE=91058534; PubMed=2244890;
RA Minegish T., Nakamura K., Takakura Y., Miyamoto K., Hasegawa Y.,
RA Ibuki Y., Igarashi M.;
RT "Cloning and sequencing of human LH/hCG receptor cDNA.";
RL Biochem. Biophys. Res. Commun. 172:1049-1054(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=92017881; PubMed=1922095;
RA Jia X.-C., Oikawa M., Bo M., Tanaka T., Ny T., Boime I., Hsueh A.J.W.;
RT "Expression of human luteinizing hormone (LH) receptor: interaction
RT with LH and chorionic gonadotropin from human but not equine, rat,
RT and ovine species.";
RL Mol. Endocrinol. 5:759-768(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Thyroid;
RX MEDLINE=91155962; PubMed=2293030;
RA Frazier A.L., Robbins L.S., Stork P.J., Sprengel R., Segaloff D.L.,
RA Cone R.D.;
RT "Isolation of TSH and LH/CG receptor cDNAs from human thyroid:
RT regulation by tissue specific splicing.";
RL Mol. Endocrinol. 4:1264-1276(1990).
RN [4]
RP 3D-STRUCTURE MODELING OF 51-232.
RX MEDLINE=96363672; PubMed=8747461;
RA Jiang X., Dreano M., Buckler D.R., Cheng S., Ythier A., Wu H.,
RA Hendrickson W.A., el Tayar N.;
RT "Structural predictions for the ligand-binding region of glycoprotein
RT hormone receptors and the nature of hormone-receptor interactions.";
RL Structure 3:1341-1353(1995).
RN [5]
RP VARIANT FMPP GLY-578.
RX MEDLINE=94019815; PubMed=7692306;
RA Shenker A., Laue L., Kosugi S., Merendino J.J. Jr., Minegishi T.,
RA Cutler G.B. Jr.;
RT "A constitutively activating mutation of the luteinizing hormone
RT receptor in familial male precocious puberty.";
RL Nature 365:652-654(1993).
RN [6]
RP VARIANTS FMPP ILE-571 AND GLY-578.
RX MEDLINE=94108425; PubMed=8281137;
RA Kremer H., Mariman E., Otten B.J., Moll G.W. Jr., Stoelinga G.B.A.,
RA Wit J.M., Jansen M., Drop S.L., Faas B., Ropers H.-H., Brunner H.G.;
RT "Cosegregation of missense mutations of the luteinizing hormone
RT receptor gene with familial male-limited precocious puberty.";
RL Hum. Mol. Genet. 2:1779-1783(1993).
RN [7]
RP VARIANT FMPP ILE-577.
RX MEDLINE=95276728; PubMed=7757065;
RA Kosugi S., van Dop C., Geffner M.E., Rabl W., Carel J.-C.,
RA Chaussain J.-L., Mori T., Merendino J.J. Jr., Shenker A.;
RT "Characterization of heterogeneous mutations causing constitutive
RT activation of the luteinizing hormone receptor in familial male
RT precocious puberty.";
RL Hum. Mol. Genet. 4:183-188(1995).
RN [8]
RP VARIANT FMPP VAL-572.
RX MEDLINE=95229804; PubMed=7714085;
RA Yano K., Saji M., Hidaka A., Moriya N., Okuno A., Kohn L.D.,
RA Cutler G.B. Jr.;
RT "A new constitutively activating point mutation in the luteinizing
RT hormone/choriogonadotropin receptor gene in cases of male-limited
RT precocious puberty.";
RL J. Clin. Endocrinol. Metab. 80:1162-1168(1995).
RN [9]
RP VARIANT FMPP VAL-568.
RX MEDLINE=95355560; PubMed=7629248;
RA Latronico A.C., Anasti J., Arnhold I.J., Mendonca B.B., Domenice S.,
RA Albano M.C., Zachman K., Wajchenberg B.L., Tsigos C.;
RT "A novel mutation of the luteinizing hormone receptor gene causing
RT male gonadotropin-independent precocious puberty.";
RL J. Clin. Endocrinol. Metab. 80:2490-2494(1995).
RN [10]
RP VARIANT LCH PRO-593.
RX MEDLINE=95235561; PubMed=7719343;
RA Kremer H., Kraaij R., Toledo S.P.A., Post M., Fridman J.B.,
RA Hayashida C.Y., van Reen M., Milgrom E., Ropers H.-H., Mariman E.,
RA Themmen A.P.N., Brunner H.G.;
RT "Male pseudohermaphroditism due to a homozygous missense mutation of
RT the luteinizing hormone receptor gene.";
RL Nat. Genet. 9:160-164(1995).
RN [11]
RP VARIANT FMPP ILE-577.
RX MEDLINE=96233936; PubMed=8829636;
RA Cocco S., Meloni A., Marini M.G., Cao A., Moi P.;
RT "A missense (T577I) mutation in the luteinizing hormone receptor gene
RT associated with familial male-limited precocious puberty.";
RL Hum. Mutat. 7:164-166(1996).
RN [12]
RP VARIANT FMPP THR-398.
RX MEDLINE=97083378; PubMed=8929952;
RA Evans B.A.J., Bowen D.J., Smith P.J., Clayton P.E., Gregory J.W.;
RT "A new point mutation in the luteinising hormone receptor gene in
RT familial and sporadic male limited precocious puberty: genotype does
RT not always correlate with phenotype.";
RL J. Med. Genet. 33:143-147(1996).
RN [13]
RP VARIANT LCH TYR-616.
RX MEDLINE=96157015; PubMed=8559204;
RA Latronico A.C., Anasti J., Arnhold I.J.P., Rapaport R., Mendonca B.B.,
RA Bloise W., Castro M., Tsigos C., Chrousos G.P.;
RT "Brief report: testicular and ovarian resistance to luteinizing
RT hormone caused by inactivating mutations of the luteinizing
RT hormone-receptor gene.";
RL New Engl. J. Med. 334:507-512(1996).
RN [14]
RP VARIANT LCH ARG-131.
RX MEDLINE=97358168; PubMed=9215288;
RA Misrahi M., Meduri G., Pissard S., Bouvattier C., Beau I.,
RA Loosfelt H., Jolivet A., Rappaport R., Milgrom E., Bougnères P.;
RT "Comparison of immunocytochemical and molecular features with the
RT phenotype in a case of incomplete male pseudohermaphroditism
RT associated with a mutation of the luteinizing hormone receptor.";
RL J. Clin. Endocrinol. Metab. 82:2159-2165(1997).
RN [15]
RP VARIANTS LEU-GLN-13 INS; SER-284 AND ASN-306.
RA Wu S.-M., Jose M., Hallermeier K., Rennert O.M., Chan W.-Y.;
RT "Polymorphisms in the coding exons of the human luteinizing hormone
RT receptor gene.";
RL Hum. Mutat. 11:333-334(1998).
RN [16]
RP VARIANT FMPP VAL-373.
RX MEDLINE=98128689; PubMed=9467560;
RA Gromoll J., Partsch C.-J., Simoni M., Nordhoff V., Sippell W.G.,
RA Nieschlag E., Saxena B.B.;
RT "A mutation in the first transmembrane domain of the lutropin receptor
RT causes male precocious puberty.";
RL J. Clin. Endocrinol. Metab. 83:476-480(1998).
RN [17]
RP VARIANT LCH LYS-354.
RX MEDLINE=98289384; PubMed=9626144;
RA Stavrou S.S., Zhu Y.S., Cai L.Q., Katz M.D., Herrera C.,
RA Defillo-Ricart M., Imperato-Mcginley J.;
RT "A novel mutation of the human luteinizing hormone receptor in 46XY
RT and 46XX sisters.";
RL J. Clin. Endocrinol. Metab. 83:2091-2098(1998).
RN [18]
RP VARIANT FMPP ARG-457.
RX MEDLINE=98326270; PubMed=9661624;
RA Latronico A.C., Abell A.N., Arnhold I.J., Liu X., Lins T.S.,
RA Brito V.N., Billerbeck A.E., Segaloff D.L., Mendonca B.B.;
RT "A unique constitutively activating mutation in third transmembrane
RT helix of luteinizing hormone receptor causes sporadic male
RT gonadotropin-independent precocious puberty.";

```
RL   J. Clin. Endocrinol. Metab. 83:2435-2440(1998).
RN   [19]
RP   VARIANT LCH 608-LEU-VAL-609 DEL.
RX   MEDLINE=98173094; PubMed=9514160;
RA   Latronico A.C., Chai Y., Arnhold I.J.P., Liu X., Mendonca B.B.,
RA   Segaloff D.L.;
RT   "A homozygous microdeletion in helix 7 of the luteinizing hormone
RT   receptor associated with familial testicular and ovarian resistance is
RT   due to both decreased cell surface expression and impaired effector
RT   activation by the cell surface receptor.";
RL   Mol. Endocrinol. 12:442-450(1998).
RN   [20]
RP   VARIANT LCH LYS-625.
RX   MEDLINE=98289955; PubMed=9626653;
RA   Martens J.W., Verhoef-Post M., Abelin N., Ezabella M., Toledo S.P.,
RA   Brunner H.G., Themmen A.P.;
RT   "A homozygous mutation in the luteinizing hormone receptor causes
RT   partial Leydig cell hypoplasia: correlation between receptor activity
RT   and phenotype.";
RL   Mol. Endocrinol. 12:775-784(1998).
RN   [21]
RP   VARIANT LEYDIG CELL TUMOR HIS-578.
RX   MEDLINE=20037768; PubMed=10580072;
RA   Liu G., Duranteau L., Carel J.-C., Monroe J., Doyle D.A., Shenker A.;
RT   "Leydig-cell tumors caused by an activating mutation of the gene
RT   encoding the luteinizing hormone receptor.";
RL   New Engl. J. Med. 341:1731-1736(1999).
CC   -!- FUNCTION: RECEPTOR FOR LUTROPIN-CHORIOGONADOTROPIC HORMONE.
CC       THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH
CC       ACTIVATE ADENYLATE CYCLASE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; A LONG FORM (SHOWN
CC       HERE) AND A SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: GONADAL AND THYROID CELLS.
CC   -!- DISEASE: DEFECTS IN LHCGR ARE A CAUSE OF FAMILIAL MALE
CC       PRECOCIOUS PUBERTY (FMPP) (ALSO KNOWN AS TESTOTOXICOSIS). IN FMPP
CC       THE RECEPTOR IS CONSTITUTIVELY ACTIVATED.
CC   -!- DISEASE: DEFECTS IN LHCGR ARE A CAUSE OF LEYDIG CELL HYPOPLASIA;
CC       (LCH) AN AUTOSOMAL RECESSIVE DISEASE CHARACTERIZED BY MALE
CC       PSEUDOHERMAPHRODITISM. IN LCH THE TESTES ARE SMALL WITH MARKED

  Query Match             25.2%;  Score 138.5;  DB 1;  Length 699;
  Best Local Similarity   27.9%;  Pred. No. 4.1e-06;
  Matches   29;  Conservative   25;  Mismatches   43;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|::|| ||   || :: :::   |  |       ::  :| :   ::|:|  |  |:
Db     534 LNVVAFFIICACYIKIYFAVRNPELMAT------NKDTKIAKKMAILIFTDFTCMAPISF 587

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
             | : |:|  |  | :  ::: | |:||  || || : |  |:
Db     588 FAISAAFKVPLITVTNSKVLLVLFYPINSCANPFLYAIFTKTFQ 631


RESULT 13
LSHR_RAT
ID   LSHR_RAT       STANDARD;      PRT;   700 AA.
AC   P16235; P70646; Q63807; Q63808; Q63809;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Lutropin-choriogonadotropic hormone receptor precursor (LH/CG-R)
DE   (LSH-R) (Luteinizing hormone receptor).
GN   LHCGR.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89332512; PubMed=2502842;
RA   McFarland K.C., Sprengel R., Phillips H.S., Koehler M.,
RA   Rosemblit N., Nikolics K., Segaloff D.L., Seeburg P.H.;
RT   "Lutropin-choriogonadotropin receptor: an unusual member of the G
RT   protein-coupled receptor family.";
RL   Science 245:494-499(1989).
RN   [2]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC   STRAIN=Sprague-Dawley; TISSUE=Ovary;
RX   MEDLINE=92347604; PubMed=1353463;
RA   Aatsinki J.T., Pietila E.M., Lakkakorpi J.T., Rajaniemi H.J.;
RT   "Expression of the LH/CG receptor gene in rat ovarian tissue is
RT   regulated by an extensive alternative splicing of the primary
RT   transcript.";
RL   Mol. Cell. Endocrinol. 84:127-135(1992).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91209270; PubMed=2019252;
RA   Koo Y.B., Slaughter R.G., Ji T.H.;
RT   "Structure of the luteinizing hormone receptor gene and multiple
RT   exons of the coding sequence.";
RL   Endocrinology 128:2297-2308(1991).
RN   [4]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX   MEDLINE=91006819; PubMed=1976554;
RA   Bernard M.P., Myers R.V., Moyle W.R.;
RT   "Cloning of rat lutropin (LH) receptor analogs lacking the soybean
RT   lectin domain.";
RL   Mol. Cell. Endocrinol. 71:R19-R23(1990).
RN   [5]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX   MEDLINE=91126285; PubMed=2281186;
RA   Segaloff D.L., Sprengel R., Nikolics K., Ascoli M.;
RT   "Structure of the lutropin/choriogonadotropin receptor.";
RL   Recent Prog. Horm. Res. 46:261-303(1990).
RN   [6]
RP   SEQUENCE OF 295-700 FROM N.A.
RX   MEDLINE=91060531; PubMed=2174034;
RA   Tsai-Morris C.H., Buczko E., Wang W., Dufau M.L.;
RT   "Intronic nature of the rat luteinizing hormone receptor gene defines
RT   a soluble receptor subspecies with hormone binding activity.";
RL   J. Biol. Chem. 265:19385-19388(1990).
RN   [7]
RP   SEQUENCE OF 27-37.
RX   MEDLINE=89174723; PubMed=2925659;
RA   Roche P.C., Ryan R.J.;
RT   "Purification, characterization, and amino-terminal sequence of rat
RT   ovarian receptor for luteinizing hormone/human choriogonadotropin.";
RL   J. Biol. Chem. 264:4636-4641(1989).
RN   [8]
RP   MUTAGENESIS.
RX   MEDLINE=91332007; PubMed=1714448;
RA   Ji I., Ji T.H.;
RT   "Asp383 in the second transmembrane domain of the lutropin receptor
RT   is important for high affinity hormone binding and cAMP production.";
RL   J. Biol. Chem. 266:14953-14957(1991).
CC   -!- FUNCTION: RECEPTOR FOR LUTROPIN-CHORIOGONADOTROPIC HORMONE.
CC       THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH
CC       ACTIVATE ADENYLATE CYCLASE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS: AT LEAST 11 ISOFORMS WHICH DIFFER IN
CC       SUBCELLULAR LOCATION ARE PRODUCED BY ALTERNATIVE SPLICING
CC       OF THE SAME GENE.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       FSH/LSH/TSH SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 7 LEUCINE-RICH REPEATS (LRR).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M26199; AAA41528.1; -.
```

```
DR   EMBL; M61212; AAA41527.1; -.
DR   EMBL; M61211; AAA41527.1; JOINED.
DR   EMBL; S40803; AAB22680.1; -.
DR   EMBL; S40787; AAB22680.1; JOINED.
DR   EMBL; S40903; AAB22680.1; JOINED.
DR   EMBL; S40904; AAB22680.1; JOINED.
DR   EMBL; S40905; AAB22680.1; JOINED.
DR   EMBL; S40907; AAB22680.1; JOINED.
DR   EMBL; S40909; AAB22680.1; JOINED.
DR   EMBL; S40918; AAB22680.1; JOINED.
DR   EMBL; S40920; AAB22680.1; JOINED.
DR   EMBL; S40795; AAB22680.1; JOINED.
DR   EMBL; S40798; AAB22680.1; JOINED.
DR   EMBL; S40795; AAB22681.1; -.
DR   EMBL; S40787; AAB22681.1; JOINED.
DR   EMBL; S40903; AAB22681.1; JOINED.
DR   EMBL; S40904; AAB22681.1; JOINED.
DR   EMBL; S40905; AAB22681.1; JOINED.
DR   EMBL; S40907; AAB22681.1; JOINED.
DR   EMBL; S40909; AAB22681.1; JOINED.
DR   EMBL; S40918; AAB22681.1; JOINED.
DR   EMBL; S40920; AAB22681.1; JOINED.
DR   EMBL; S40803; AAB22682.2; -.
DR   EMBL; S40787; AAB22682.2; JOINED.
DR   EMBL; S40903; AAB22682.2; JOINED.
DR   EMBL; S40907; AAB22682.2; JOINED.
DR   EMBL; S40909; AAB22682.2; JOINED.
DR   EMBL; S40918; AAB22682.2; JOINED.
DR   EMBL; S40920; AAB22682.2; JOINED.
DR   EMBL; S40795; AAB22682.2; JOINED.
DR   EMBL; S40798; AAB22682.2; JOINED.
DR   EMBL; S40803; AAB22683.1; -.
DR   EMBL; S40787; AAB22683.1; JOINED.
DR   EMBL; S40903; AAB22683.1; JOINED.
DR   EMBL; S40904; AAB22683.1; JOINED.
DR   EMBL; S40905; AAB22683.1; JOINED.
DR   EMBL; S40907; AAB22683.1; JOINED.
DR   EMBL; S40909; AAB22683.1; JOINED.
DR   EMBL; S40918; AAB22683.1; JOINED.
DR   EMBL; S40920; AAB22683.1; JOINED.
DR   EMBL; S40795; AAB22683.1; JOINED.
DR   EMBL; S40798; AAB22683.1; JOINED.
DR   EMBL; S40803; AAB22684.2; -.
DR   EMBL; S40787; AAB22684.2; JOINED.
DR   EMBL; S40903; AAB22684.2; JOINED.
DR   EMBL; S40904; AAB22684.2; JOINED.
DR   EMBL; S40905; AAB22684.2; JOINED.
DR   EMBL; S40909; AAB22684.2; JOINED.
DR   EMBL; S40918; AAB22684.2; JOINED.
DR   EMBL; S40920; AAB22684.2; JOINED.
DR   EMBL; S40795; AAB22684.2; JOINED.
DR   EMBL; S40798; AAB22684.2; JOINED.
DR   EMBL; M68928; AAA41529.1; -.
DR   EMBL; M68917; AAA41529.1; JOINED.
DR   EMBL; M68918; AAA41529.1; JOINED.
DR   EMBL; M68919; AAA41529.1; JOINED.
DR   EMBL; M68920; AAA41529.1; JOINED.
DR   EMBL; M68921; AAA41529.1; JOINED.
DR   EMBL; M68922; AAA41529.1; JOINED.
DR   EMBL; M68923; AAA41529.1; JOINED.
DR   EMBL; M68925; AAA41529.1; JOINED.
DR   EMBL; M68926; AAA41529.1; JOINED.
DR   EMBL; M68927; AAA41529.1; JOINED.
DR   PIR; A32460; A32460.
DR   PIR; A41343; A41343.
DR   HSSP; P22888; 1LUT.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 1.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing.
FT   SIGNAL       1     26
FT   CHAIN       27    700       LUTROPIN-CHORIOGONADOTROPIC HORMONE
FT                               RECEPTOR.
FT   DOMAIN       27    362       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    363    390       1 (POTENTIAL).
FT   DOMAIN      391    399       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    400    422       2 (POTENTIAL).
FT   DOMAIN      423    443       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    444    466       3 (POTENTIAL).
FT   DOMAIN      467    486       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    487    509       4 (POTENTIAL).
FT   DOMAIN      510    529       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    530    551       5 (POTENTIAL).
FT   DOMAIN      552    574       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    575    598       6 (POTENTIAL).
FT   DOMAIN      599    609       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    610    631       7 (POTENTIAL).
FT   DOMAIN      632    700       CYTOPLASMIC (POTENTIAL).
FT   REPEAT       52     75       LRR 1.
FT   REPEAT      126    150       LRR 2.
FT   REPEAT      152    175       LRR 3.
FT   REPEAT      176    200       LRR 4.
FT   REPEAT      202    224       LRR 5.
FT   REPEAT      225    248       LRR 6.
FT   REPEAT      250    271       LRR 7.
FT   DISULFID    443    518       BY SIMILARITY.
FT   CARBOHYD    103    103       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    178    178       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    199    199       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    295    295       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    303    303       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    317    317       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC     83    132       MISSING (IN ISOFORM 1950).
FT   VARSPLIC    133    157       MISSING (IN ISOFORM 1759).

  Query Match             24.9%;  Score 136.5;  DB 1;  Length 700;
  Best Local Similarity   27.9%;  Pred. No. 6.4e-06;
  Matches   29;  Conservative   25;  Mismatches   43;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|::||::|   || :: ::|   |          ::  :| :   ::|:|  |  |:
Db     538 LNVVAFVVICACYIRIYFAVQNPELTAP------NKDTKIAKKMAILIFTDFTCMAPISF 591

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
             | : |:|  |  | :  ::: | ||||  || || : |  |:
Db     592 FAISAAFKVPLITVTNSKILLVLFYPVNSCANPFLYAIFTKAFQ 635


RESULT 14
FSHR_CHICK
ID   FSHR_CHICK     STANDARD;      PRT;   693 AA.
AC   P79763; Q90719;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Follicle stimulating hormone receptor precursor (FSH-R) (Follitropin
DE   receptor).
GN   FSHR.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=97473503; PubMed=9332357;
RA   Wakabayashi N., Suzuki A., Hoshino H., Nishimori K., Mizuno S.;
RT   "The cDNA cloning and transient expression of a chicken gene encoding
RT   a follicle-stimulating hormone receptor.";
```

```
RL   Gene 197:121-127(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=White leghorn;
RX   MEDLINE=97057887; PubMed=8902217;
RA   You S., Bridgham J.T., Foster D.N., Johnson A.L.;
RT   "Characterization of the chicken follicle-stimulating hormone
RT   receptor (cFSH-R) complementary deoxyribonucleic acid, and expression
RT   of cFSH-R messenger ribonucleic acid in the ovary.";
RL   Biol. Reprod. 55:1055-1062(1996).
CC   -!- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
CC       OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC       ADENYLATE CYCLASE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       FSH/LSH/TSH SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 7 LEUCINE-RICH REPEATS (LRR).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D87871; BAA13487.1; -.
DR   EMBL; U51097; AAC60030.1; -.
DR   HSSP; P23945; 1XUN.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Phosphorylation; Repeat; Leucine-rich repeat.
FT   SIGNAL       1     17       POTENTIAL.
FT   CHAIN       18    693       FOLLICLE STIMULATING HORMONE RECEPTOR.
FT   DOMAIN      18    366       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   367    387       1 (POTENTIAL).
FT   DOMAIN     388    398       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   399    421       2 (POTENTIAL).
FT   DOMAIN     422    443       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   444    465       3 (POTENTIAL).
FT   DOMAIN     466    485       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   486    508       4 (POTENTIAL).
FT   DOMAIN     509    528       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   529    550       5 (POTENTIAL).
FT   DOMAIN     551    573       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   574    597       6 (POTENTIAL).
FT   DOMAIN     598    608       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   609    630       7 (POTENTIAL).
FT   DOMAIN     631    693       CYTOPLASMIC (POTENTIAL).
FT   REPEAT      45     68       LRR 1.
FT   REPEAT      69     93       LRR 2.
FT   REPEAT      95    118       LRR 3.
FT   REPEAT     119    143       LRR 4.
FT   REPEAT     168    192       LRR 5.
FT   REPEAT     193    216       LRR 6.
FT   REPEAT     218    240       LRR 7.
FT   DISULFID   442    517       BY SIMILARITY.
FT   CARBOHYD    47     47       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   191    191       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   199    199       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   268    268       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     4      4       G -> D (IN REF. 2).
FT   CONFLICT    88     88       G -> A (IN REF. 2).
FT   CONFLICT   140    140       K -> R (IN REF. 2).
FT   CONFLICT   174    174       I -> T (IN REF. 2).
FT   CONFLICT   191    191       N -> S (IN REF. 2).
FT   CONFLICT   329    329       V -> L (IN REF. 2).
SQ   SEQUENCE   693 AA;  78697 MW;  46F98699635A1BEC CRC64;

  Query Match             24.1%;  Score 132.5;  DB 1;  Length 693;
  Best Local Similarity   27.9%;  Pred. No. 1.5e-05;
  Matches   29;  Conservative   25;  Mismatches   43;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|:|||:||   || :: :::   : ::       :  :| |   ::|:| :|  |:
Db     537 LNVLAFVIICICYICIYFTVRNPNVISS------NSDTKIAKRMAILIFTDFLCMAPISF 590

Qy      62 VKILSLFRVEIPDTMTSWI-VIFFLPVNSALNPILYTLTTNFFK 104
             | :  || :     | | :: | |:||  || || : |  |:
Db     591 FAISASLRVPLITVSKSKILLVLFYPINSCANPFLYAIFTKTFR 634


RESULT 15
LSHR_MOUSE
ID   LSHR_MOUSE     STANDARD;      PRT;   700 AA.
AC   P30730;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Lutropin-choriogonadotropic hormone receptor precursor (LH/CG-R)
DE   (LSH-R) (Luteinizing hormone receptor).
GN   LHCGR OR LHR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92165799; PubMed=1311310;
RA   Gudermann T., Birnbaumer M., Birnbaumer L.;
RT   "Evidence for dual coupling of the murine luteinizing hormone
RT   receptor to adenylyl cyclase and phosphoinositide breakdown and Ca2+
RT   mobilization. Studies with the cloned murine luteinizing hormone
RT   receptor expressed in L cells.";
RL   J. Biol. Chem. 267:4479-4488(1992).
RN   [2]
RP   SEQUENCE OF 1-58 FROM N.A.
RX   MEDLINE=93093308; PubMed=1459341;
RA   Huhtaniemi I.T., Eskola V., Pakarinen P., Matikainen T., Sprengel R.;
RT   "The murine luteinizing hormone and follicle-stimulating hormone
RT   receptor genes: transcription initiation sites, putative promoter
RT   sequences and promoter activity.";
RL   Mol. Cell. Endocrinol. 88:55-66(1992).
CC   -!- FUNCTION: RECEPTOR FOR LUTROPIN-CHORIOGONADOTROPIC HORMONE.
CC       THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH
CC       ACTIVATE ADENYLATE CYCLASE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       FSH/LSH/TSH SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 4 LEUCINE-RICH REPEATS (LRR).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S49753; AAB24402.1; -.
DR   EMBL; M81310; AAA39432.1; -.
DR   EMBL; M87571; AAA39433.1; -.
DR   PIR; A42395; A42395.
DR   HSSP; P22888; 1LUT.
DR   MGD; MGI:96783; Lhcgr.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
```

```
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 1.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Phosphorylation; Repeat; Leucine-rich repeat.
FT   SIGNAL        1     26       BY SIMILARITY.
FT   CHAIN        27    700       LUTROPIN-CHORIOGONADOTROPIC HORMONE
FT                                 RECEPTOR.
FT   DOMAIN        27    362       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     363    390       1 (POTENTIAL).
FT   DOMAIN       391    399       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     400    422       2 (POTENTIAL).
FT   DOMAIN       423    443       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     444    466       3 (POTENTIAL).
FT   DOMAIN       467    486       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     487    509       4 (POTENTIAL).
FT   DOMAIN       510    529       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     530    551       5 (POTENTIAL).
FT   DOMAIN       552    574       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     575    598       6 (POTENTIAL).
FT   DOMAIN       599    609       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     610    631       7 (POTENTIAL).
FT   DOMAIN       632    700       CYTOPLASMIC (POTENTIAL).
FT   REPEAT        52     75       LRR 1.
FT   REPEAT       126    150       LRR 2.
FT   REPEAT       176    200       LRR 3.
FT   REPEAT       225    248       LRR 4.
FT   DISULFID     443    518       BY SIMILARITY.
FT   CARBOHYD     103    103       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     178    178       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     199    199       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     295    295       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     303    303       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     317    317       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   700 AA;  78214 MW;  8A6840A011E1E014 CRC64;

  Query Match             24.1%;  Score 132.5;  DB 1;  Length 700;
  Best Local Similarity   26.9%;  Pred. No. 1.5e-05;
  Matches   28;  Conservative   25;  Mismatches   44;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :| :||::|   |: :: ::|   |          ::  :| :   ::|:|  |  |:
Db     538 LNAVAFVVICACYVRIYFAVQNPELTAP------NKDTKIAKKMAILIFTDFTCMAPISF 591

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
             | : |:|  |  | :  ::: | ||||  || || : |  |:
Db     592 FAISAAFKVPLITVTNSKVLLVLFYPVNSCANPFLYAVFTKAFQ 635


Search completed: May 22, 2003, 12:45:32
Job time : 26 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: May 22, 2003, 12:42:05 ; Search time 84 Seconds
(without alignments)
262.465 Million cell updates/sec

Title: US-09-930-312-2
Perfect score: 549
Sequence: 1 GVNLLAFLIIVFSYITMFCS.....NSALNPILYTLTTNFFKDKL 107

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	234	42.6	334	5	Q9VYG0	Q9vyg0 drosophila
2	225	41.0	359	5	Q9VBP0	Q9vbp0 drosophila
3	158	28.8	1012	5	Q95YI6	Q95yi6 asterina pe
4	158	28.8	1280	5	Q95YI7	Q95yi7 asterina pe
5	145.5	26.5	701	6	Q8SPS8	Q8sps8 bos taurus
6	142	25.9	1050	5	Q9BN18	Q9bn18 drosophila
7	142	25.9	1300	5	Q9NKD6	Q9nkd6 drosophila
8	142	25.9	1360	5	Q9NDI1	Q9ndi1 drosophila
9	138.5	25.2	601	13	O42500	O42500 meleagris g
10	138.5	25.2	693	13	Q9DGC5	Q9dgc5 oreochromis
11	138.5	25.2	699	4	Q15996	Q15996 homo sapien
12	138.5	25.2	701	4	Q14751	Q14751 homo sapien
13	136.5	24.9	335	6	O46387	O46387 mustela vis
14	135.5	24.7	814	13	Q9I949	Q9i949 oncorhynchu
15	134.5	24.5	793	13	Q9I948	Q9i948 oncorhynchu
16	132.5	24.1	658	13	Q9PVN9	Q9pvn9 oncorhynchu
17	131.5	24.0	293	6	Q95MF7	Q95mf7 canis famil
18	131.5	24.0	724	13	Q9PVP0	Q9pvp0 oncorhynchu
19	128.5	23.4	779	13	Q9I8N7	Q9i8n7 morone saxa
20	127.5	23.2	696	13	Q98T85	Q98t85 ictalurus p
21	126.5	23.0	763	6	Q9BGN4	Q9bgn4 felis silve
22	126.5	23.0	778	13	Q98TF4	Q98tf4 oreochromis
23	126	23.0	829	5	Q9VEG4	Q9veg4 drosophila
24	126	23.0	831	5	Q94979	Q94979 drosophila
25	126	23.0	831	5	Q8SX01	Q8sx01 drosophila
26	125.5	22.9	696	13	Q9DGF5	Q9dgf5 cynops pyrr
27	125	22.8	410	4	Q16225	Q16225 homo sapien
28	124	22.6	459	11	Q8R301	Q8r301 mus musculu
29	123	22.4	326	4	Q8TAM8	Q8tam8 homo sapien
30	123	22.4	695	11	Q8R428	Q8r428 cavia porce
31	123	22.4	701	13	Q9DGC6	Q9dgc6 oreochromis
32	122	22.2	688	11	Q64183	Q64183 rattus sp.
33	121	22.0	326	11	O08766	O08766 rattus norv
34	120	21.9	310	11	Q8R0M5	Q8r0m5 mus musculu
35	120	21.9	928	4	Q9BYD7	Q9byd7 homo sapien
36	119.5	21.8	470	4	O43200	O43200 homo sapien
37	119.5	21.8	533	4	Q9P1V4	Q9p1v4 homo sapien
38	118.5	21.6	319	6	Q9N1U0	Q9n1u0 oryctolagus
39	118.5	21.6	662	13	Q9PW16	Q9pw16 clarias gar
40	118.5	21.6	764	11	Q9D697	Q9d697 mus musculu
41	117.5	21.4	739	6	Q9BG55	Q9bg55 sus scrofa
42	117.5	21.4	764	6	Q9BG56	Q9bg56 sus scrofa
43	117.5	21.4	764	6	Q8SPP9	Q8spp9 sus scrofa
44	116.5	21.2	662	13	Q98T84	Q98t84 ictalurus p
45	116	21.1	107	11	Q99P16	Q99p16 mus musculu

ALIGNMENTS

RESULT 1
Q9VYG0
ID Q9VYG0 PRELIMINARY; PRT; 334 AA.
AC Q9VYG0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CG4187 protein.
GN CG4187.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

```
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003491; AAF48237.1; -.
DR   FlyBase; FBgn0030458; CG4187.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001901; SecE.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS01067; SECE_SEC61G; UNKNOWN_1.
SQ   SEQUENCE   334 AA;  37020 MW;  1662F2D467534061 CRC64;

  Query Match             42.6%;  Score 234;  DB 5;  Length 334;
  Best Local Similarity   45.4%;  Pred. No. 1.4e-13;
  Matches   49;  Conservative   20;  Mismatches   33;  Indels    6;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKT--ALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPV 59
           || |: : |:|||| |  :|: :   :::|      |||  || ||  || :|  ||:|:
Db     174 VNTLSLIFILFSYIRMLQAIRDSGGGMRSTH----SGRENVVATRFAIIVTTDCACWLPI 229

Qy      60 FVVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
           |||: :|   ||   : :|: :  ||||||||||:||||||  || :|
Db     230 IVVKLAALSGCEISPDLYAWLAVLVLPVNSALNPVLYTLTTAAFKQQL 277


RESULT 2
Q9VBP0
ID   Q9VBP0      PRELIMINARY;      PRT;   359 AA.
AC   Q9VBP0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   CG5042 protein.
GN   CG5042 OR CG5046.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AE003753; AAF56490.1; -.
DR   FlyBase; FBgn0039354; CG5042.
DR   InterPro; IPR002106; AAtRNA_ligaseII.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Glycoprotein; Transmembrane.
SQ   SEQUENCE   359 AA;  40880 MW;  71913BEBD7C9E739 CRC64;

  Query Match             41.0%;  Score 225;  DB 5;  Length 359;
  Best Local Similarity   42.6%;  Pred. No. 1.5e-17;
  Matches   46;  Conservative   23;  Mismatches   33;  Indels    6;  Gaps    3;

Qy       1 GVNLLAFLIIVFSYITMFCSIQKTALQT-TEVRNCFGREVAVANRFFFIVFSDAICWIPV 59
             |||||  ::|   |  :  || :|   |   : :|   | ||  |||||| :| :||:|:
Db     207 GVNLLLLVMIAMLYTALLISIWRTRSATPLTLLDC---EFAV--RFFFIVLTDFLCWVPI 261

Qy      60 FVVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKL 107
            |:||   |   | | : :|:|:| ||:|||:||:||| ||  :::::
Db     262 IVMKIWVFFNYNISDDIYAWLVVFVLPLNSAVNPLLYTFTTPKYRNQI 309


RESULT 3
Q95YI6
ID   Q95YI6      PRELIMINARY;      PRT;  1012 AA.
AC   Q95YI6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Glycoprotein hormone receptor.
GN   APGPHR'.
OS   Asterina pectinifera (Starfish).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
OC   Asteroidea; Valvatacea; Valvatida; Asterinidae; Asterina.
OX   NCBI_TaxID=7594;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Mita M., Hirai T., Oba Y., Yoshikuni M., Nagahama Y.;
RT   "cDNA cloning and functional analysis of a novel member of the
RT   glycoprotein hormone receptor family from a starfish Asterina
RT   pectinifera.";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB061862; BAB68209.1; -.
DR   InterPro; IPR002106; AAtRNA_ligaseII.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
```

```
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 14.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   1012 AA;  112623 MW;  52A70E7A88C46E0A CRC64;

  Query Match             28.8%;  Score 158;  DB 5;  Length 1012;
  Best Local Similarity   33.3%;  Pred. No. 1.5e-09;
  Matches   37;  Conservative   22;  Mismatches   30;  Indels   22;  Gaps    4;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPV-- 59
           :|:|||:||:  | :::      |:| :   ||   :  || |   :||:|  || |:
Db     729 LNILAFVIIMACYASIY-----LAIQGSHAWNC--NDSRVARRMSLLVFTDFACWAPIAF 781

Qy      60 ------FVVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
                 | ::::||   ::       : || ||:||  || |||: |  ||
Db     782 FSLTAAFGLRLISLDGAKV-------LTIFVLPLNSCANPFLYTILTKQFK 825


RESULT 4
Q95YI7
ID   Q95YI7      PRELIMINARY;      PRT;  1280 AA.
AC   Q95YI7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Glycoprotein hormone receptor.
GN   APGPHR.
OS   Asterina pectinifera (Starfish).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
OC   Asteroidea; Valvatacea; Valvatida; Asterinidae; Asterina.
OX   NCBI_TaxID=7594;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Mita M., Hirai T., Oba Y., Yoshikuni M., Nagahama Y.;
RT   "cDNA cloning and functional analysis of a novel member of the
RT   glycoprotein hormone receptor family from a starfish Asterina
RT   pectinifera.";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB061861; BAB68208.1; -.
DR   InterPro; IPR002106; AAtRNA_ligaseII.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 14.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   1280 AA;  141700 MW;  0AB0ECCC0DD880BA CRC64;

  Query Match             28.8%;  Score 158;  DB 5;  Length 1280;
  Best Local Similarity   33.3%;  Pred. No. 1.9e-09;
  Matches   37;  Conservative   22;  Mismatches   30;  Indels   22;  Gaps    4;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPV-- 59
           :|:|||:||:  | :::      |:| :   ||   :  || |   :||:|  || |:
Db     729 LNILAFVIIMACYASIY-----LAIQGSHAWNC--NDSRVARRMSLLVFTDFACWAPIAF 781

Qy      60 ------FVVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
                 | ::::||   ::       : || ||:||  || |||: |  ||
Db     782 FSLTAAFGLRLISLDGAKV-------LTIFVLPLNSCANPFLYTILTKQFK 825


RESULT 5
Q8SPS8
ID   Q8SPS8      PRELIMINARY;      PRT;  701 AA.
AC   Q8SPS8;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Luteinizing hormone receptor.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HOLSTEIN; TISSUE=CORPUS LUTEUM;
RA   Kawate N., Tamada H., Inaba T., Sawada T.;
RT   "Molecular cloning of a full-length cDNA encoding bovine luteinizing
RT   hormone receptor and its expression in COS-7 cells.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF491303; AAM09535.1; -.
KW   Receptor.
SQ   SEQUENCE   701 AA;  78455 MW;  D70AE862EB265CCF CRC64;

  Query Match             26.5%;  Score 145.5;  DB 6;  Length 701;
  Best Local Similarity   29.8%;  Pred. No. 2.9e-08;
  Matches   31;  Conservative   24;  Mismatches   42;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|::||:||   || :: ::|   |  |       ::  :| :   ::|:|  |  |:
Db     536 LNVMAFIIICACYIKIYFAVQNPELMAT------NKDTKIAKKMAVLIFTDFTCMAPISF 589

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
             | : |:|  |  | :  ::: | ||||  || || : |  |:
Db     590 FAISAAFKVPLITVTNSKVLLVLFYPVNSCANPFLYAIFTKAFQ 633


RESULT 6
Q9BN18
ID   Q9BN18      PRELIMINARY;      PRT;  1050 AA.
AC   Q9BN18;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Leucine-rich repeat-containing G protein-coupled receptor 2.
GN   RK OR CG8930.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Nishi S., Hsu S.Y., Zell K., Hsueh A.J.;
RT   "Characterization of two fly LGR (leucine-rich repeat-containing G
RT   protein-coupled receptor) proteins homologous to vertebrate
RT   glycoprotein hormone receptors: constitutively activation of wild type
RT   fly LGR1 but not LGR2 in transfected mammalian cells.";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF274591; AAK00808.1; -.
DR   FlyBase; FBgn0003255; rk.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 10.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   SMART; SM00370; LRR; 8.
DR   SMART; SM00369; LRR_TYP; 10.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   1050 AA;  117707 MW;  35D71260A8B4BF99 CRC64;

  Query Match             25.9%;  Score 142;  DB 5;  Length 1050;
```

```
  Best Local Similarity   31.4%;  Pred. No. 1e-07;
  Matches   33;  Conservative   22;  Mismatches   40;  Indels   10;  Gaps    3;

Qy       2 VNLLAFLIIVFSYITMFCSIQ-KTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
           :|  ||| ::  |: |: :|:   |  | : |       :| |   :||:| :|| |:
Db     715 INGCAFLTLMGCYLKMYWAIRGSQAWNTNDSR--------IAKRMALLVFTDFLCWSPIA 766

Qy      61 VVKILSLFRVEIPDTMTSWI-VIFFLPVNSALNPILYTLTTNFFK 104
             | ::| :::     : |  :| ||:||  || || : |  ||
Db     767 FFSITAIFGLQLISLEQAKIFTVFVLPLNSCCNPFLYAIMTKQFK 811


RESULT 7
Q9NKD6
ID   Q9NKD6      PRELIMINARY;      PRT;  1300 AA.
AC   Q9NKD6; Q9VJU3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical 144.0 kDa protein (Rk gene product).
GN   RK OR CG8930.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=99403001; PubMed=10471707;
RA   Ashburner M., Misra S., Roote J., Lewis S.E., Blazej R., Davis T.,
RA   Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
RA   Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
RA   Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitelaw K.,
RA   Celniker S., Rubin G.M.;
RT   "An exploration of the sequence of a 2.9-Mb region of the genome of
RT   Drosophila melanogaster: the Adh region.";
RL   Genetics 153:179-219(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RA   Celniker S.E., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G.,
RA   Butenhoff C., Champe M., Chavez C., Ciesiolka L., Doyle C.M.,
RA   Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
RA   Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim E., Lee B.,
RA   Lewis S., Li P., Lomotan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
RA   Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,
RA   Sethi H., Snir E., Svirskas R.R., Wan K.H., Weinburg T., Zhang R.,
RA   Zieran L.L., Rubin G.M.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003408; AAF44846.1; -.
DR   EMBL; AE003642; AAF53367.2; -.
DR   FlyBase; FBgn0003255; rk.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 10.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   SMART; SM00370; LRR; 1.
DR   SMART; SM00369; LRR_TYP; 4.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   1300 AA;  144031 MW;  B4B9E39F942FA0B3 CRC64;

  Query Match           25.9%;  Score 142;  DB 5;  Length 1300;
  Best Local Similarity   31.4%;  Pred. No. 1.2e-07;
  Matches   33;  Conservative   22;  Mismatches   40;  Indels   10;  Gaps    3;

Qy       2 VNLLAFLIIVFSYITMFCSIQ-KTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
           :|  ||| ::  |: |: :|:   |  | : |       :| |   :||:E :|| |:
Db     869 INGCAFLTLMGCYLKMYWAIRGSQAWNTNDSR--------IAKRMALLVFTDFLCWSPIA 920

Qy      61 VVKILSLFRVEIPDTMTSWI-VIFFLPVNSALNPILYTLTTNFFK 104
             | ::| :::     : |  :| ||:||  || || : |  ||
Db     921 FFSITAIFGLQLISLEQAKIFTVFVLPLNSCCNPFLYAIMTKQFK 965


RESULT 8
Q9NDI1
ID   Q9NDI1      PRELIMINARY;      PRT;  1360 AA.
AC   Q9NDI1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Glycoprotein hormone receptor II.
GN   RK OR CG8930.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CANTON S; TISSUE=WHOLE ANIMAL;
RX   MEDLINE=20359836; PubMed=10899142;
RA   Eriksen K.K., Hauser F., Schiott M., Pedersen K.-M., Soendergaard L.,
RA   Grimmelikhuijzen C.J.P.;
RT   "Molecular Cloning, Genomic Organization, Developmental Regulation,
RT   and a Knock-Out Mutant of a Novel Leu-Rich Repeats-Containing G
RT   Protein-Coupled Receptor (DLGR-2) from Drosophila melanogaster.";
RL   Genome Res. 10:924-938(2000).
DR   EMBL; AF142343; AAF66608.1; -.
```

```
DR   HSSP; Q57815; 1D3Y.
DR   FlyBase; FBgn0003255; rk.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR003592; LRR_out.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 14.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   SMART; SM00370; LRR; 2.
DR   SMART; SM00369; LRR_TYP; 5.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   1360 AA;  150731 MW;  7D435155B4F6F612 CRC64;

  Query Match             25.9%;  Score 142;  DB 5;  Length 1360;
  Best Local Similarity   31.4%;  Pred. No. 1.3e-07;
  Matches   33;  Conservative   22;  Mismatches   40;  Indels   10;  Gaps    3;

Qy       2 VNLLAFLIIVFSYITMFCSIQ-KTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVF 60
           :|  ||| ::  |: |: :|:   |  | : |        :| |   :||:| :|| |:
Db     929 INGCAFLTLMGCYLKMYWAIRGSQAWNTNDSR--------IAKRMALLVFTDFLCWSPIA 980

Qy      61 VVKILSLFRVEIPDTMTSWI-VIFFLPVNSALNPILYTLTTNFFK 104
             | ::| :::     : |  :| ||:||  || || : |  ||
Db     981 FFSITAIFGLQLISLEQAKIFTVFVLPLNSCCNPFLYAIMTKQFK 1025


RESULT 9
O42500
ID   O42500      PRELIMINARY;      PRT;   601 AA.
AC   O42500;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Luteinizing hormone receptor (Fragment).
OS   Meleagris gallopavo (Common turkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.
OX   NCBI_TaxID=9103;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20078926; PubMed=10611074;
RA   You S., Kim H., Hsu C.C., El Halawani M.E., Foster D.N.;
RT   "Three different turkey luteinizing hormone receptor (tLH-R) isoforms
RT   I: characterization of alternatively spliced tLH-R isoforms and their
RT   regulated expression in diverse tissues.";
RL   Biol. Reprod. 62:108-116(2000).
DR   EMBL; U92082; AAB64409.1; -.
DR   HSSP; P22888; 1LUT.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
SQ   SEQUENCE   601 AA;  67380 MW;  714D2FD6D5E28FDC CRC64;

  Query Match             25.2%;  Score 138.5;  DB 13;  Length 601;
  Best Local Similarity   28.8%;  Pred. No. 1.6e-07;
  Matches   30;  Conservative   23;  Mismatches   44;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :| :||::|   || :: ::|   |         ::  :| |   ::|:|  |  |:
Db     432 LNAVAFVVICACYIKIYIAVQNPELVAA------NKDTKIAKRMAILIFTDFTCMAPISF 485

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
             | : |:|  |  | :  ::: | ||||  || || : |  |:
Db     486 FAISAAFKVPLITVTNSKILLVLFYPVNSCANPFLYAIFTKAFQ 529


RESULT 10
```

```
Q9DGC5
ID   Q9DGC5      PRELIMINARY;      PRT;   693 AA.
AC   Q9DGC5;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Gonadotropin receptor II.
GN   TGTH-RII.
OS   Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC   Cichlidae; Oreochromis.
OX   NCBI_TaxID=8128;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Oba Y., Hirai T., Yoshiura Y., Yao Z., Nagahama Y.;
RT   "Tilapia gonadotropin receptor II.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB041763; BAB16107.1; -.
DR   HSSP; P22888; 1LUT.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 2.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   693 AA;  77856 MW;  C3BBEFD8ECFC8988 CRC64;

  Query Match             25.2%;  Score 138.5;  DB 13;  Length 693;
  Best Local Similarity   28.8%;  Pred. No. 1.8e-07;
  Matches   30;  Conservative   23;  Mismatches   44;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|::|||:: : || :: |:       |       :  :| |   ::|:| :|  |:
Db     534 LNVVAFLVVCYCYICIYLSVHNPEHSTRR------GDTKIAKRMAVLIFTDFLCMAPISF 587

Qy      62 VKILSLFRVEIPDTMTSWI-VIFFLPVNSALNPILYTLTTNFFK 104
             | :  |: :     | | :| | |:||  || |||: |  |:
Db     588 FAISAALRMPLITVSHSKILLILFYPINSLCNPFLYTIFTRAFR 631


RESULT 11
Q15996
ID   Q15996      PRELIMINARY;      PRT;   699 AA.
AC   Q15996;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Luteinizing hormone receptor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OVARY;
RX   MEDLINE=92017881; PubMed=1922095;
RA   Jia X.C., Oikawa M., Bo M., Tanaka T., Ny T., Boime I., Hsueh A.J.;
RT   "Expression of human luteinizing hormone (LH) receptor: interaction
RT   with LH and chorionic gonadotropin from human but not equine, rat, and
RT   ovine species.";
RL   Mol. Endocrinol. 5:759-768(1991).
DR   EMBL; S57793; AAB19917.2; -.
DR   HSSP; P22888; 1LUT.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 3.
DR   SMART; SM00013; LRRNT; 1.
```

```
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   699 AA;  78642 MW;  2E3D93F4621BA842 CRC64;

  Query Match             25.2%;  Score 138.5;  DB 4;  Length 699;
  Best Local Similarity   27.9%;  Pred. No. 1.8e-07;
  Matches   29;  Conservative   25;  Mismatches   43;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|::|| ||   || :: :::   |  |       ::  :| :   ::|:|  |  |:
Db     534 LNVVAFFIICACYIKIYFAVRNPELMAT------NKDTKIAKKMAILIFTDFTCMAPISF 587

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
             | : |:|  |  | :  ::: | |:||  || || : |  |:
Db     588 FAISAAFKVPLITVTNSKVLLVLFYPINSCANPFLYAIFTKTFQ 631


RESULT 12
Q14751
ID   Q14751      PRELIMINARY;      PRT;   701 AA.
AC   Q14751;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   LUTEINISING hormone-choriogonadotropin receptor (Luteinizing hormone
DE   receptor).
GN   LHCGR OR LHR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96023956; PubMed=7556872;
RA   Atger M., Misrahi M., Sar S., Leflem L., Dessen P., Milgrom E.;
RT   "Structure of the human luteinizing hormone-choriogonadotropin
RT   receptor gene: unusual promoter and 5' non-coding regions.";
RL   Mol. Cell. Endocrinol. 111:113-123(1995).
RN   [2]
RP   SEQUENCE OF 1-56 FROM N.A.
RA   Tsai-Morris C.H., Geng Y., Buczko E., Dufau M.L.;
RT   "Human luteinizing hormone receptor gene containing 6 bp insertion in
RT   exon 1 coding region with identical 5' flanking sequence as LHR2 exon
RT   1 coding region.";
RL   Hum. Hered. 49:0-0(1999).
DR   EMBL; X84753; CAA59234.1; -.
DR   EMBL; X84754; CAA59234.1; JOINED.
DR   EMBL; X84755; CAA59234.1; JOINED.
DR   EMBL; X84756; CAA59234.1; JOINED.
DR   EMBL; X84757; CAA59234.1; JOINED.
DR   EMBL; X84758; CAA59234.1; JOINED.
DR   EMBL; X84759; CAA59234.1; JOINED.
DR   EMBL; X84760; CAA59234.1; JOINED.
DR   EMBL; X84761; CAA59234.1; JOINED.
DR   EMBL; X84762; CAA59234.1; JOINED.
DR   EMBL; X84763; CAA59234.1; JOINED.
DR   EMBL; AF082076; AAC98291.1; -.
DR   HSSP; P22888; 1LUT.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 3.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   701 AA;  78757 MW;  E69FFF5BABC409D4 CRC64;

  Query Match             25.2%;  Score 138.5;  DB 4;  Length 701;
  Best Local Similarity   27.9%;  Pred. No. 1.8e-07;
  Matches   29;  Conservative   25;  Mismatches   43;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|::|| ||   || :: :::   |  |       ::  :| :   ::|:|  |  |:
Db     536 LNVVAFFIICACYIKIYFAVRNPELMAT------NKDTKIAKKMAILIFTDFTCMAPISF 589

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
             | : |:|  |  | :  ::: | |:||  || || : |  |:
Db     590 FAISAAFKVPLITVTNSKVLLVLFYPINSCANPFLYAIFTKTFQ 633


RESULT 13
O46387
ID   O46387      PRELIMINARY;      PRT;   335 AA.
AC   O46387;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Luteinizing hormone receptor (Fragment).
GN   LHR.
OS   Mustela vison (American mink).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Mustelidae; Mustelinae;
OC   Mustela.
OX   NCBI_TaxID=9667;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS;
RX   MEDLINE=98384190; PubMed=9716555;
RA   Douglas D.A., Houde A., Song J.H., Farookhi R., Concannon P.W.,
RA   Murphy B.D.;
RT   "Luteotropic hormone receptors in the ovary of the mink (Mustela
RT   vison) during delayed implantation and early-postimplantation
RT   gestation.";
RL   Biol. Reprod. 59:571-578(1998).
DR   EMBL; AF029295; AAB88900.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
FT   NON_TER       1      1
SQ   SEQUENCE   335 AA;  37880 MW;  7C647BB7BEF2C1C2 CRC64;

  Query Match             24.9%;  Score 136.5;  DB 6;  Length 335;
  Best Local Similarity   28.8%;  Pred. No. 1.6e-07;
  Matches   30;  Conservative   23;  Mismatches   44;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFV 61
           :|::|| ||   || :: ::|   |          ::  :| :   ::|:|  |  |:
Db     174 LNVVAFTIICACYIKIYFAVQNPELMAN------NKDTKIAKKMAVLIFTDFTCMAPISF 227

Qy      62 VKILSLFRVE-IPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
             | : |:|  |  | :  ::: | ||||  || || : |  |:
Db     228 FAISAAFKVPLITVTNSKVLLVLFYPVNSCANPFLYAIFTKAFQ 271


RESULT 14
Q9I949
ID   Q9I949      PRELIMINARY;      PRT;   814 AA.
AC   Q9I949;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Thyrotropin receptor A.
GN   STSH-RA.
OS   Oncorhynchus rhodurus (amago).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=41164;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RX   MEDLINE=20462948; PubMed=11006115;
RA   Oba Y., Hirai T., Yoshiura Y., Kobayashi T., Nagahama Y.;
RT   "Cloning, Functional Characterization, and Expression of Thyrotropin
RT   Receptors in the Thyroid of Amago Salmon (Oncorhynchus rhodurus).";
RL   Biochem. Biophys. Res. Commun. 276:258-263(2000).
DR   EMBL; AB030954; BAB07800.1; -.
DR   HSSP; P16473; 1XUM.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   814 AA;  91091 MW;  1FC9A2AE8D6C911E CRC64;

  Query Match             24.7%;  Score 135.5;  DB 13;  Length 814;
  Best Local Similarity   27.9%;  Pred. No. 4.5e-07;
  Matches   29;  Conservative   25;  Mismatches   43;  Indels    7;  Gaps    2;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPV-F 60
           :|:||||:|   || ::|::     ::  '      ::  :| |   ::|:| :|  |: |
Db     634 LNILAFLVICACYIKIYCTVHNPHYRSG------SKDTNIAKRMAVLIFTDFLCMAPISF 687

Qy      61 VVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
                | |  |  : :  ::: | |:||  || || : |  |:
Db     688 YAMSAVLDRPLITVSNSKILLVLFYPLNSCANPFLYAIFTKAFR 731


RESULT 15
Q9I948
ID   Q9I948      PRELIMINARY;      PRT;   793 AA.
AC   Q9I948;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Thyrotropin receptor B.
GN   STSH-RB.
OS   Oncorhynchus rhodurus (amago).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=41164;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20462948; PubMed=11006115;
RA   Oba Y., Hirai T., Yoshiura Y., Kobayashi T., Nagahama Y.;
RT   "Cloning, Functional Characterization, and Expression of Thyrotropin
RT   Receptors in the Thyroid of Amago Salmon (Oncorhynchus rhodurus).";
RL   Biochem. Biophys. Res. Commun. 276:258-263(2000).
DR   EMBL; AB030955; BAB07801.1; -.
DR   HSSP; P16473; 1XUM.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   793 AA;  89410 MW;  2842BF0D67ED93DC CRC64;

  Query Match             24.5%;  Score 134.5;  DB 13;  Length 793;
  Best Local Similarity   27.6%;  Pred. No. 5.7e-07;
  Matches   29;  Conservative   25;  Mismatches   42;  Indels    9;  Gaps    3;

Qy       2 VNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFG-REVAVANRFFFIVFSDAICWIPV- 59
           :|:||||:|   |: ::|::            | | ::  :| |   ::|:| :|  |:
Db     608 LNILAFLVICTCYVKIYCAVHNPYY-----CSGSKDTNIAKRMAILIFTDFLCMAPIS 660

Qy      60 FVVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFK 104
           |      : |  |  : :  ::: | |:||  || || : |  |:
Db     661 FYAMSAVVDRPLITVSNSKILLVLFYPLNSCANPFLYAIFTKAFR 705


Search completed: May 22, 2003, 12:47:04
Job time : 86 secs
```

GenCore version 5.1.4_p5_4578

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 26, 2003, 19:35:57 ; Search time 2362 Seconds
(without alignments)
1318.374 Million cell updates/sec

Title: US-09-930-312-2
Perfect score: 549
Sequence: 1 GVNLLAFLIIVFSYITMFCS.....NSALNPILYTLTTNFFKDKL 107

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09930312/runat_22052003_114613_27065/app_query.fasta_1.263
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09930312_@CGN_1_1_2496_@runat_22052003_114613_27065 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	549	100.0	321	6	AX384516	AX384516 Sequence
	2	549	100.0	1068	6	AX148176	AX148176 Sequence
	3	549	100.0	2049	6	AX385037	AX385037 Sequence
	4	549	100.0	2142	6	AX451566	AX451566 Sequence
	5	549	100.0	2193	6	AX385032	AX385032 Sequence
	6	549	100.0	2214	6	AX451562	AX451562 Sequence
	7	549	100.0	2265	6	AX385027	AX385027 Sequence
	8	549	100.0	2436	9	AF453828	AF453828 Homo sapi
	9	549	100.0	2838	9	AF403384	AF403384 Homo sapi
	10	478	87.1	2214	6	AX385045	AX385045 Sequence
	11	478	87.1	2539	10	AF346501	AF346501 Mus muscu
	12	376	68.5	1191	6	AX128507	AX128507 Sequence
	13	376	68.5	2274	9	AF190500	AF190500 Homo sapi
c	14	371	67.6	1018	6	AX147820	AX147820 Sequence
	15	371	67.6	170522	9	AL138708	AL138708 Human DNA
	16	349	63.6	474	6	AX088165	AX088165 Sequence
c	17	326.5	59.5	164610	2	AC098990	AC098990 Rattus no
	18	326.5	59.5	171869	2	AC121029	AC121029 Rattus no
c	19	319	58.1	180673	10	AC068627	AC068627 Mus Muscu
c	20	319	58.1	213462	10	AC077689	AC077689 Mus Muscu
	21	255	46.4	187431	9	AC107219	AC107219 Homo sapi
	22	253.5	46.2	58377	2	AC128964	AC128964 Rattus no
c	23	253.5	46.2	124282	2	AC121415	AC121415 Rattus no
	24	253.5	46.2	173546	2	AC098607	AC098607 Rattus no
c	25	251.5	45.8	179055	2	AC122462	AC122462 Mus muscu
	26	234	42.6	1015	6	AX254453	AX254453 Sequence
	27	225	41.0	1080	6	AX254459	AX254459 Sequence
	28	225	41.0	3187	6	AX254458	AX254458 Sequence
c	29	225	41.0	57274	2	AC017540	AC017540 Drosophil
	30	225	41.0	163970	3	AC010580	AC010580 Drosophil
	31	225	41.0	167396	3	AC010579	AC010579 Drosophil
	32	225	41.0	226332	3	AE003753	AE003753 Drosophil
	33	210	38.3	3756	3	LSGPCR	Z23104 L.stagnalis
	34	188	34.2	176565	2	AL136106	AL136106 Homo sapi
	35	181	33.0	3574	6	AX254452	AX254452 Sequence
c	36	181	33.0	15570	2	AC014247	AC014247 Drosophil
c	37	181	33.0	171751	3	AC023697	AC023697 Drosophil
c	38	181	33.0	174157	3	AC023711	AC023711 Drosophil
c	39	181	33.0	323461	3	AE003491	AE003491 Drosophil
	40	158.5	28.9	137956	2	AC010918	AC010918 Drosophil
	41	158	28.8	3942	3	AB061862	AB061862 Asterina
	42	158	28.8	3943	3	AB061861	AB061861 Asterina
c	43	156	28.4	171869	2	AC121029	AC121029 Rattus no
c	44	154	28.1	173546	2	AC098607	AC098607 Rattus no
	45	146.5	26.7	5886	10	AF069778	AF069778 Mus muscu

ALIGNMENTS

RESULT 1

```
AX384516
LOCUS       AX384516                321 bp    DNA     linear   PAT 19-MAR-2002
DEFINITION  Sequence 1 from Patent WO0214496.
ACCESSION   AX384516
VERSION     AX384516.1  GI:19577719
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lind,P.
  TITLE     Novel g protein-coupled receptors
  JOURNAL   Patent: WO 0214496-A 1 21-FEB-2002;
            PHARMACIA & UPJOHN COMPANY (US)
FEATURES             Location/Qualifiers
     source          1. .321
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       76 a     67 c     59 g    119 t
ORIGIN

Alignment Scores:
Pred. No.:              2.98e-51        Length:         321
Score:                  549.00          Matches:        107
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-930-312-2 (1-107) x AX384516 (1-321)

Qy      1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60

Qy     21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120

Qy     41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180

Qy     61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240

Qy     81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300

Qy    101 AsnPhePheLysAspLysLeu 107
          |||||||||||||||||||||
Db    301 AACTTTTTTAAGGACAAGTTG 321

RESULT 2
AX148176
LOCUS       AX148176                1068 bp    DNA     linear   PAT 08-JUN-2001
DEFINITION  Sequence 17 from Patent WO0136471.
ACCESSION   AX148176
VERSION     AX148176.1  GI:14347081
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1068)
  AUTHORS   Chen,R., Dang,H.T. and Lowitz,K.P.
  TITLE     Endogenous and non-endogenous versions of human g protein-coupled
            receptors
  JOURNAL   Patent: WO 0136471-A 17 25-MAY-2001;
            Arena Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1068
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      278 a    215 c    214 g    361 t
ORIGIN

Alignment Scores:
Pred. No.:              9.26e-51        Length:         1068
Score:                  549.00          Matches:        107
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-930-312-2 (1-107) x AX148176 (1-1068)

Qy      1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    589 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 648

Qy     21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 708

Qy     41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 768

Qy     61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 828

Qy     81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    829 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 888

Qy    101 AsnPhePheLysAspLysLeu 107
          |||||||||||||||||||||
Db    889 AACTTTTTTAAGGACAAGTTG 909

RESULT 3
AX385037
LOCUS       AX385037                2049 bp    DNA     linear   PAT 19-MAR-2002
DEFINITION  Sequence 11 from Patent WO0214489.
ACCESSION   AX385037
VERSION     AX385037.1  GI:19578160
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Paszty,C.J., Gong,J., Daugherty,B. and Rogers,N.
  TITLE     Leucine-rich repeat-containing g-protein coupled receptor-8
            molecules and uses thereof
  JOURNAL   Patent: WO 0214489-A 11 21-FEB-2002;
            Amgen, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2049
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             1. .2049
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD28590.1"
                     /db_xref="GI:19578161"
                     /translation="MIVFLVFKHLFSLRLITMFFLLHFIVLINVKDFALTQGSMITPS
                     CQKGYFPCGNLTKCLPRAFHCDGKDDCGNGADEENCGDTSGWATIFGTVHGNANSVAL
                     TQECFLKQYPQCCDCKETELECVNGDLKSVPMISNNVTLLSLKKNKIHSLPDKVFIKY
                     TKLKKIFLQHNCIRHISRKAFFGLCNLQILILDDNPITRISQRLFTGLNSLFFLSMVN
                     NYLEALPKQMCAQMPQLNWVDLEGNRIKYLTNSTFLSCDSLTVLDLSSNTITELSPHL
                     FKDLKLLQKLDLERIEIPNINTRMFQPMKNLSHIYFKNFRYCSYAPHVRICMPLTDGI
```

SSFEDLLANNILRIFVWVIAFITCFGNLFVIGMRSFIKAENTTHAMSIKILCCADCLM
GVYLFFVGIFDIKYRGQYQKYALLWMESVQCRLMGFLAMLSTEVSVLLLTYLTLEKFL
VIVFPFSNIRPGKRQTSVILICIWMAGFLIAVIPFWNKDYFGNFYGKNGVCFPLYYDQ
TEDIGSKGYSLGIFLGVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVAN
RFFFIVFSDAICWIPVFVVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTN
FFKDKLKQLLHKHQRKSIFKIKKKSLSTSIVWIEDSSSLKLGVLNKITLGDSIMKPVS
"

sig_peptide 1. .108
BASE COUNT 579 a 415 c 395 g 660 t
ORIGIN

Alignment Scores:
Pred. No.: 1.71e-50 Length: 2049
Score: 549.00 Matches: 107
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-930-312-2 (1-107) x AX385037 (1-2049)

```
Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1570 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1629

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1630 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1689

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1690 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1749

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1750 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1809

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1810 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 1869

Qy     101 AsnPhePheLysAspLysLeu 107
          |||||||||||||||||||||
Db    1870 AACTTTTTTAAGGACAAGTTG 1890
```

RESULT 4
AX451566
LOCUS AX451566 2142 bp DNA linear PAT 03-JUL-2002
DEFINITION Sequence 5 from Patent WO0226824.
ACCESSION AX451566
VERSION AX451566.1 GI:21698550
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Feder,J.N., Mintier,G., Ramanathan,C.S. and Hawken,D.R.
TITLE A novel human g-protein coupled receptor, hgprbmy5, expressed highly in brain and ovarian tissues
JOURNAL Patent: WO 0226824-A 5 04-APR-2002;
Bristol-Myers Squibb Company (US)
FEATURES Location/Qualifiers
source 1. .2142
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 613 a 435 c 408 g 686 t
ORIGIN

Alignment Scores:
Pred. No.: 1.79e-50 Length: 2142
Score: 549.00 Matches: 107
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-930-312-2 (1-107) x AX451566 (1-2142)

```
Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1663 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1722

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1723 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1782

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1783 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1842

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1843 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1902

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1903 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 1962

Qy     101 AsnPhePheLysAspLysLeu 107
          |||||||||||||||||||||
Db    1963 AACTTTTTTAAGGACAAGTTG 1983
```

RESULT 5
AX385032
LOCUS AX385032 2193 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 6 from Patent WO0214489.
ACCESSION AX385032
VERSION AX385032.1 GI:19578156
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Paszty,C.J., Gong,J., Daugherty,B. and Rogers,N.
TITLE Leucine-rich repeat-containing g-protein coupled receptor-8 molecules and uses thereof
JOURNAL Patent: WO 0214489-A 6 21-FEB-2002;
Amgen, Inc. (US)
FEATURES Location/Qualifiers
source 1. .2193
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS 1. .2193
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD28588.1"
/db_xref="GI:19578157"
/translation="MIVFLVFKHLFSLRLITMFFLLHFIVLINVKDFALTQGSMITPS
CQKGYFPCGNLTKCLPRAFHCDGKDDCGNGADEENCGDTSGWATIFGTVHGNANSVAL
TQECFLKQYPQCCDCKETELECVNGDLKSVPMISNNVTLLSLKKNKIHSLPDKVFIKY
TKLKKIFLQHNCIRHISRKAFFGLCNLQILYLNHNCITTLRPGIFKDLHQLTWLILDD
NPITRISQRLFTGLNSLFFLSMVNNYLEALPKQMCAQMPQLNWVDLEGNRIKYLTNST
FLSCDSLTVLDLSSNTITELSPHLFKDLKLLQKLNLSSNPLMYLHKNQFESLKQLQSL
DLERIEIPNINTRMFQPMKNLSHIYFKNFRYCSYAPHVRICMPLTDGISSFEDLLANN
ILRIFVWVIAFITCFGNLFVIGMRSFIKAENTTHAMSIKILCCADCLMGVYLFFVGIF
DIKYRGQYQKYALLWMESVQCRLMGFLAMLSTEVSVLLLTYLTLEKFLVIVFPFSNIR
PGKRQTSVILICIWMAGFLIAVIPFWNKDYFGNFYGKNGVCFPLYYDQTEDIGSKGYS
LGIFLGVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDA
ICWIPVFVVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDKLKQLL
HKHQRKSIFKIKKKSLSTSIVWIEDSSSLKLGVLNKITLGDSIMKPVS"
sig_peptide 1. .108
BASE COUNT 626 a 456 c 411 g 700 t
ORIGIN

```
Alignment Scores:
Pred. No.:                1.83e-50        Length:          2193
Score:                    549.00          Matches:         107
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       6               Gaps:            0

US-09-930-312-2 (1-107) x AX385032 (1-2193)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1714 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1773

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1774 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1833

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1834 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1893

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1894 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1953

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1954 GTGATTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2013

Qy     101 AsnPhePheLysAspLysLeu 107
           |||||||||||||||||||||
Db    2014 AACTTTTTTAAGGACAAGTTG 2034

RESULT 6
AX451562
LOCUS       AX451562               2214 bp    DNA     linear   PAT 03-JUL-2002
DEFINITION  Sequence 1 from Patent WO0226824.
ACCESSION   AX451562
VERSION     AX451562.1  GI:21698547
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Feder,J.N., Mintier,G., Ramanathan,C.S. and Hawken,D.R.
  TITLE     A novel human g-protein coupled receptor, hgprbmy5, expressed
            highly in brain and ovarian tissues
  JOURNAL   Patent: WO 0226824-A 1 04-APR-2002;
            Bristol-Myers Squibb Company (US)
FEATURES             Location/Qualifiers
     source          1. .2214
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      638 a    457 c    416 g    703 t
ORIGIN

Alignment Scores:
Pred. No.:                1.84e-50        Length:          2214
Score:                    549.00          Matches:         107
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       6               Gaps:            0

US-09-930-312-2 (1-107) x AX451562 (1-2214)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1735 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1794

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1795 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1854

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1855 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1914

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1915 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1974

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1975 GTGATTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2034

Qy     101 AsnPhePheLysAspLysLeu 107
           |||||||||||||||||||||
Db    2035 AACTTTTTTAAGGACAAGTTG 2055

RESULT 7
AX385027
LOCUS       AX385027               2265 bp    DNA     linear   PAT 19-MAR-2002
DEFINITION  Sequence 1 from Patent WO0214489.
ACCESSION   AX385027
VERSION     AX385027.1  GI:19578152
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Paszty,C.J., Gong,J., Daugherty,B. and Rogers,N.
  TITLE     Leucine-rich repeat-containing g-protein coupled receptor-8
            molecules and uses thereof
  JOURNAL   Patent: WO 0214489-A 1 21-FEB-2002;
            Amgen, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2265
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             1. .2265
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD28586.1"
                     /db_xref="GI:19578153"
                     /translation="MIVFLVFKHLFSLRLITMFFLLHFIVLINVKDFALTQGSMITPS
                     CQKGYFPCGNLTKCLPRAFHCDGKDDCGNGADEENCGDTSGWATIFGTVHGNANSVAL
                     TQECFLKQYPQCCDCKETELECVNGDLKSVPMISNNVTLLSLKKNKIHSLPDKVFIKY
                     TKLKKIFLQHNCIRHISRKAFFGLCNLQILYLNHNCITTLRPGIFKDLHQLTWLILDD
                     NPITRISQRLFTGLNSLFFLSMVNNYLEALPKQMCAQMPQLNWVDLEGNRIKYLTNST
                     FLSCDSLTVLFLPRNQIGFVPEKTFSSLKNLGELDLSSNTITELSPHLFKDLKLLQKL
                     NLSSNPLMYLHKNQFESLKQLQSLDLERIEIPNINTRMFQPMKNLSHIYFKNFRYCSY
                     APHVRICMPLTDGISSFEDLLANNILRIFVWVIAFITCFGNLFVIGMRSFIKAENTTH
                     AMSIKILCCADCLMGVYLFFVGIFDIKYRGQYQKYALLWMESVQCRLMGFLAMLSTEV
                     SVLLLTYLTLEKFLVIVFPFSNIRPGKRQTSVILICIWMAGFLIAVIPFWNKDYFGNF
                     YGKNGVCFPLYYDQTEDIGSKGYSLGIFLGVNLLAFLIIVFSYITMFCSIQKTALQTT
                     EVRNCFGREVAVANRFFFIVFSDAICWIPVFVVKILSLFRVEIPDTMTSWIVIFFLPV
                     NSALNPILYTLTTNFFKDKLKQLLHKHQRKSIFKIKKKSLSTSIVWIEDSSSLKLGVL
                     NKITLGDSIMKPVS"
     sig_peptide     1. .108
BASE COUNT      650 a    466 c    423 g    726 t
ORIGIN

Alignment Scores:
Pred. No.:                1.88e-50        Length:          2265
Score:                    549.00          Matches:         107
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       6               Gaps:            0

US-09-930-312-2 (1-107) x AX385027 (1-2265)
```

```
Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
Db    1786 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1845

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
Db    1846 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1905

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
Db    1906 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1965

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
Db    1966 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 2025

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
Db    2026 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2085

Qy     101 AsnPhePheLysAspLysLeu 107
Db    2086 AACTTTTTTAAGGACAAGTTG 2106

RESULT 8
AF453828
LOCUS       AF453828               2436 bp    mRNA    linear   PRI 04-FEB-2002
DEFINITION  Homo sapiens G protein-coupled receptor affecting testicular
            descent (GREAT) mRNA, complete cds.
ACCESSION   AF453828
VERSION     AF453828.1  GI:18483167
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2436)
  AUTHORS   Gorlov,I.P., Kamat,A., Jones,E., Lamb,D., Truong,A., Bogatcheva,N.,
            Bishop,C.E., McElreavey,K. and Agoulnik,A.I.
  TITLE     Mutations of the GREAT gene cause cryptorchidism
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2436)
  AUTHORS   Agoulnik,A.I.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-NOV-2001) Ob/Gyn, Baylor College of Medicine, 6550
            Fannin St., Su. 861, Houston, TX 77030, USA
FEATURES             Location/Qualifiers
     source          1. .2436
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="13"
                     /map="13q12-q13"
     gene            1. .2436
                     /gene="GREAT"
     CDS             71. .2335
                     /gene="GREAT"
                     /codon_start=1
                     /product="G protein-coupled receptor affecting testicular
                     descent"
                     /protein_id="AAL73946.1"
                     /db_xref="GI:18483168"
                     /translation="MIVFLVFKHLFSLRLITMFFLLHFIVLINVKDFALTQGSMITPS
                     CQKGYFPCGNLTKCLPRAFHCDGKDDCGNGADEENCGDTSGWATIFGTVHGNANSVAL
                     TQECFLKQYPQCCDCKETELECVNGDLKSVPMISNNVTLLSLKKNKIHSLPDKVFIKY
                     TKLKKIFLQHNCIRHISRKAFFGLCNLQILYLNHNCITTLRPGIFKDLHQLTWLILDD
                     NPITRISQRLFTGLNSLFFLSMVNNYLEALPKQMCAQMPQLNWVDLEGNRIKYLTNST
                     FLSCDSLTVLFLPRNQIGFVPEKTFSSLKNLGELDLSSNTITELSPHLFKDLKLLQKL
                     NLSSNPLMYLHKNQFESLKQLQSLDLERIEIPNINTRMFQPMKNLSHIYFKNFRYCSY
                     APHVRICMPLTDGISSFEDLLANNILRIFVWVIAFITCFGNLFVIGMRSFIKAENTTH
                     AMSIKILCCADCLMGVYLFFVGIFDIKYRGQYQKYALLWMESVQCRLMGFLAMLSTEV
                     SVLLLTYLTLEKFLVIVFPFSNIRPGKRQTSVILICIWMAGFLIAVIPFWNKDYFGNF
                     YGKNGVCFPLYYDQTEDIGSKGYSLGIFLGVNLLAFLIIVFSYITMFCSIQKTALQTT
                     EVRNCFGREVAVANRFFFIVFSDAICWIPVFVVKILSLFRVEIPDTMTSWIVIFFLPV
                     NSALNPILYTLTTNFFKDKLKQLLHKHQRKSIFKIKKKSLSTSIVWIEDSSSLKLGVL
                     NKITLGDSIMKPVS"
BASE COUNT      703 a    503 c    456 g    774 t
ORIGIN

Alignment Scores:
Pred. No.:                2.02e-50       Length:          2436
Score:                    549.00         Matches:         107
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       9              Gaps:            0

US-09-930-312-2 (1-107) x AF453828 (1-2436)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
Db    1856 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1915

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
Db    1916 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1975

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
Db    1976 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 2035

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
Db    2036 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 2095

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
Db    2096 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2155

Qy     101 AsnPhePheLysAspLysLeu 107
Db    2156 AACTTTTTTAAGGACAAGTTG 2176

RESULT 9
AF403384
LOCUS       AF403384               2838 bp    mRNA    linear   PRI 19-FEB-2002
DEFINITION  Homo sapiens LGR8 mRNA, complete cds.
ACCESSION   AF403384
VERSION     AF403384.2  GI:18702459
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2838)
  AUTHORS   Hsu,S.Y., Nakabayashi,K., Nishi,S., Kumagai,J., Kudo,M.,
            Sherwood,O.D. and Hsueh,A.J.
  TITLE     Activation of orphan receptors by the hormone relaxin
  JOURNAL   Science 295 (5555), 671-674 (2002)
  MEDLINE   21669315
   PUBMED   11809971
REFERENCE   2  (bases 1 to 2838)
  AUTHORS   Hsu,S.Y., Nakabayashi,K. and Bhalla,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUL-2001) GYN/OB, Stanford University, MSOB S385,
            Stanford, CA 94305, USA
REFERENCE   3  (bases 1 to 2838)
  AUTHORS   Hsueh,A.J.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-FEB-2002) Division of Reproductive Biology,
            Department of Gynecology and Obstetrics, Stanford University
            Medical Center, 300 Pasteur Drive, Room A-344, Stanford, CA
            94305-5317, USA
  REMARK    Sequence update by submitter
COMMENT     On Feb 19, 2002 this sequence version replaced gi:18419431.
FEATURES             Location/Qualifiers
     source          1. .2838
```

```
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11q13"
     CDS             107. .2371
                     /codon_start=1
                     /product="LGR8"
                     /protein_id="AAL69324.2"
                     /db_xref="GI:18702460"
                     /translation="MIVFLVFKHLFSLRLITMFFLLHFIVLINVKDFALTQGSMITPS
                     CQKGYFPCGNLTKCLPRAFHCDGKDDCGNGADEENCGDTSGWATIFGTVHGNANSVAL
                     TQECFLKQYPQCCDCKETELECVNGDLKSVPMISNNVTLLSLKKNKIHSLPDKVFIKY
                     TKLKKIFLQHNCIRHISRKAFFGLCNLQILYLNHNCITTLRPGIFKDLHQLTWLILDD
                     NPITRISQRLFTGLNSLFFLSMVNNYLEALPKQMCAQMPQLNWVDLEGNRIKYLTNST
                     FLSCDSLTVLFLPRNQIGFVPEKTFSSLKNLGELDLSSNTITELSPHLFKDLKLLQKL
                     NLSSNPLMYLHKNQFESLKQLQSLDLERIEIPNINTRMFQPMKNLSHIYFKNFRYCSY
                     APHVRICMPLTDGISSFEDLLANNILRIFVWVIAFITCFGNLFVIGMRSFIKAENTTH
                     AMSIKILCCADCLMGVYLFFVGIFDIKYRGQYQKYALLWMESVQCRLMGFLAMLSTEV
                     SVLLLTYLTLEKFLVIVFPFSNIRPGKRQTSVILICIWMAGFLIAVIPFWNKDYFGNF
                     YGKNGVCFPLYYDQTEDIGSKGYSLGIFLGVNLLAFLIIVFSYITMFCSIQKTALQTT
                     EVRNCFGREVAVANRFFFIVFSDAICWIPVFVVKILSLFRVEIPDTMTSWIVIFFLPV
                     NSALNPILYTLTTNFFKDKLKQLLHKHQRKSIFKIKKKSLSTSIVWIEDSSSLKLGVL
                     NKITLGDSIMKPVS"
BASE COUNT      817 a    616 c    536 g    869 t
ORIGIN

Alignment Scores:
Pred. No.:              2.33e-50        Length:         2838
Score:                  549.00          Matches:        107
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-930-312-2 (1-107) x AF403384 (1-2838)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1892 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1951

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1952 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 2011

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2012 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 2071

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2072 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 2131

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2132 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2191

Qy     101 AsnPhePheLysAspLysLeu 107
           |||||||||||||||||||||
Db    2192 AACTTTTTTAAGGACAAGTTG 2212

RESULT 10
AX385045
LOCUS       AX385045               2214 bp    DNA     linear   PAT 19-MAR-2002
DEFINITION  Sequence 19 from Patent WO0214489.
ACCESSION   AX385045
VERSION     AX385045.1  GI:19578166
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Paszty,C.J., Gong,J., Daugherty,B. and Rogers,N.
  TITLE     Leucine-rich repeat-containing g-protein coupled receptor-8
            molecules and uses thereof
  JOURNAL   Patent: WO 0214489-A 19 21-FEB-2002;
            Amgen, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2214
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     CDS             1. .2214
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD28593.1"
                     /db_xref="GI:19578167"
                     /translation="MWLLLHVILLTEVKDFALADSSMVAPLCPKGYFPCGNLTKCLPR
                     AFHCDGVDDCGNGADEDNCGDTSGWTTIFGTVHGNVNKVTLTQECFLSQYPQHCYCRE
                     NELECVKADLKAVPKVSSNVTLLSLKKNKIHRLPVKVFSRYTELRKIYLQHNCITHIS
                     RRAFLGLHNLQILYLSHNCITSLRPGIFKDLHQLAWLILDDNPITRISQKSFMGLNSL
                     FFLSMVGNRLEALPETLCAQMPQLNWVDLANNGIKYITNSTFLTCDSLTVLFLPRNQI
                     GFVPEKTFSSLKNLGELDLSSNMITKLPVHLFSDLHLLQKLNLSSNPLLYVHKNQFGS
                     LKQLQSLDLERIEIPNISTGMFQPMKNLSHIYLKTFRYCSYVPHVRICMPSTDGISSS
                     EDLLANGILRVSVWVIAFITCVGNFLVIAVRSLIKAENTTHAMSIKILCCADCLMGVY
                     LFSVGVFDIKYRGQYQKYALLWMESVPCRLLGFLATLSTEVSVLLLTFLTLEKFLVIV
                     FPFSNLRLGKRQTAVALASIWVVGFLIAAVPFTREDYFGNFYGKNGVCFPLHYDQAED
                     FGSRGYSLGIFLGVNLLAFLVIVISYVTMFCSIHKTALQTAEVRSHIGKEVAVANRFF
                     FIVFSDAICWIPVFVVKILSLLQVEIPGTITSWIVVFFLPVNSALNPILYTLTTSFFK
                     DKLKQLLHKHRRKPIFKVKKKSLSASIVWTDESSLKLGVLSKIALGDSIMKPVSP"
     sig_peptide     1. .57
BASE COUNT      554 a    548 c    492 g    620 t
ORIGIN

Alignment Scores:
Pred. No.:              9.93e-43        Length:         2214
Score:                  478.00          Matches:        92
Percent Similarity:     93.46%          Conservative:   8
Best Local Similarity:  85.98%          Mismatches:     7
Query Match:            87.07%          Indels:         0
DB:                     6               Gaps:           0

US-09-930-312-2 (1-107) x AX385045 (1-2214)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           ||||||||||||||||||||||||:::||||||   ||||||:::|||||||||||||||||
Db    1735 GGTGTGAACTTGCTGGCTTTCCTCGTCATCGTGATTTCCTATGTCACCATGTTCTGCTCC 1794

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           |||   ||||||||||||||||||   |||||||:::      |||:::||||||||||
Db    1795 ATTCATAAAACAGCCCTTCAGACTGCAGAAGTGAGGAGCCACATCGGGAAGGAGGTGGCT 1854

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1855 GTTGCAAACCGGTTCTTTTTTATCGTGTTCTCTGATGCCATCTGCTGGATCCCTGTGTTT 1914

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           ||||||||||||||||||||   :::||||||||||||   |||:::||||||||||||
Db    1915 GTCGTTAAGATCCTGTCTCTCCTTCAAGTGGAGATACCAGGCACAATCACTTCCTGGATC 1974

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           |||:::||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1975 GTGGTTTTTTTCCTTCCGGTGAACAGCGCCTTAAACCCCATCCTCTACACTCTGACGACC 2034

Qy     101 AsnPhePheLysAspLysLeu 107
           :::|||||||||||||||||||
Db    2035 TCCTTTTTTAAGGACAAGTTG 2055

RESULT 11
AF346501
LOCUS       AF346501               2539 bp    mRNA    linear   ROD 01-OCT-2001
DEFINITION  Mus musculus G protein coupled receptor affecting testicular
            descent (Great) mRNA, complete cds.
ACCESSION   AF346501
VERSION     AF346501.1  GI:15811372
KEYWORDS    .
SOURCE      Mus musculus.
```

```
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2539)
  AUTHORS   Overbeek,P.A., Gorlov,I.P., Sutherland,R.W., Houston,J.B.,
            Harrison,W.R., Boettger-Tong,H.L., Bishop,C.E. and Agoulnik,A.I.
  TITLE     A transgenic insertion causing cryptorchidism in mice
  JOURNAL   Genesis 30 (1), 26-35 (2001)
  MEDLINE   21250990
   PUBMED   11353515
REFERENCE   2  (bases 1 to 2539)
  AUTHORS   Agoulnik,A.I.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-FEB-2001) Ob/Gyn, Baylor College of Medicine, 6550
            Fannin Str. Su. 861, Houston, TX 77030, USA
FEATURES             Location/Qualifiers
     source          1. .2539
                     /organism="Mus musculus"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /chromosome="5"
     gene            1. .2539
                     /gene="Great"
     CDS             194. .2407
                     /gene="Great"
                     /note="GREAT"
                     /codon_start=1
                     /product="G protein coupled receptor affecting testicular
                     descent"
                     /protein_id="AAL08943.1"
                     /db_xref="GI:15811373"
                     /translation="MWLLLHVILLTEVKDFALADSSMVAPLCPKGYFPCGNLTKCLPR
                     AFHCDGVDDCGNGADEDNCGDTSGWTTIFGTVHGNVNKVTLTQECFLSQYPQHCYCRE
                     NELECVKADLKAVPKVSSNVTLLSLKKNKIHRLPVKVFSRYTELRKIYLQHNCITHIS
                     RRAFLGLHNLQILYLSHNCITSLRPGIFKDLHQLAWLILDDNPITRISQKSFMGLNSL
                     FFLPMVGNRLEALPETLCAQMPQLNWVDLANNGIKYITNSTFLTCDSLTVLFLPRNQI
                     GFVPEKTFSSLKNLGELDLSSNMITKLPVHLFSDLHLLQKLNLSSNPLLYVHKNQFGS
                     LKQLQSLDLERIEIPNISTGMFQPMKNLSHIYLKTFRYCSYVPHVRICMPSTDGISSS
                     EDLLANGILRVSVWVIAFITCVGNFLVIAVRSLIKAENTTHAMSIKILCCADCLMGVY
                     LFSVGVFDIKYRGQYQKYALLWMESVPCRLLGFLATLSTEVSVLLLTFLTLEKFLVIV
                     FPFSNLRLGKRQTAVALASIWVVGFLIAAVPFTREDYFGNFYGKNGVCFPLHYDQAED
                     FGSRGYSLGIFLGVNLLAFLVIVISYVTMFCSIHKTALQTAEVRSHIGKEVAVANRFF
                     FIVFSDAICWIPVFVVKILSLLQVEIPGTITSWIVVFFLPVNSALNPILYTLTTSFFK
                     DKLKQLLHKHRRKPIFKVKKKSLSASIVWTDESSLKLGVLSKIALGDSIMKPVSP"
BASE COUNT      639 a    641 c    574 g    685 t
ORIGIN

Alignment Scores:
Pred. No.:                1.13e-42       Length:          2539
Score:                    478.00         Matches:         92
Percent Similarity:       93.46%         Conservative:    8
Best Local Similarity:    85.98%         Mismatches:      7
Query Match:              87.07%         Indels:          0
DB:                       10             Gaps:            0

US-09-930-312-2 (1-107) x AF346501 (1-2539)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           ||||||||||||||||||||||||:::|||||||   |||||||:::|||||||||||||||
Db    1928 GGTGTGAACTTGCTGGCTTTCCTCGTCATCGTGATTTCCTATGTCACCATGTTCTGCTCC 1987

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           |||   |||||||||||||||||   |||||||||:::       |||:::|||||||||
Db    1988 ATTCATAAAACAGCCCTTCAGACTGCAGAAGTGAGGAGCCACATCGGGAAGGAGGTGGCT 2047

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2048 GTTGCAAACCGGTTCTTTTTTATCGTGTTCTCTGATGCCATCTGCTGGATCCCTGTGTTT 2107

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           ||||||||||||||||||||   :::||||||||||||   |||:::||||||||||||||
Db    2108 GTCGTTAAGATCCTGTCTCTCCTTCAAGTGGAGATACCAGGCACAATCACTTCCTGGATC 2167

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           |||:::||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2168 GTGGTTTTTTTCCTTCCGGTGAACAGCGCCTTAAACCCCATCCTCTACACTCTGACGACC 2227

Qy     101 AsnPhePheLysAspLysLeu 107
           :::||||||||||||||||||
Db    2228 TCCTTTTTTAAGGACAAGTTG 2248

RESULT 12
AX128507
LOCUS       AX128507               1191 bp    DNA     linear   PAT 15-MAY-2001
DEFINITION  Sequence 15 from Patent WO0131014.
ACCESSION   AX128507
VERSION     AX128507.1  GI:14134974
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1191)
  AUTHORS   Vogeli,G., Wood,L.S. and Merchant,K.
  TITLE     G protein-coupled receptors expressed in brain
  JOURNAL   Patent: WO 0131014-A 15 03-MAY-2001;
            PHARMACIA & UPJOHN COMPANY (US)
FEATURES             Location/Qualifiers
     source          1. .1191
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             1. .1191
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC38938.1"
                     /db_xref="GI:14134975"
                     /translation="MFRPLVNLSHIYFKKFQYCGYAPHVRSCKPNTDGISSLENLLAS
                     IIQRVFVWVVSAVTCFGNIFVICMRPYIRSENKLYAMSIISLCCADCLMGIYLFVIGG
                     FDLKFRGEYNKHAQLWMESTHCQLVGSLAILSTEVSVLLLTFLTLEKYICIVYPFRCV
                     RPGKCRTITVLILIWITGFIVAFIPLSNKEFFKNYYGTNGVCFPLHSEDTESIGAQIY
                     SVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVFTD
                     ALCWIPIFVVKFLSLLQVEIPGTITSWVVIFILPINSALNPILYTLTTRPFKEMIHRF
                     WYNYRQRKSMDSKGQKTYAPSFIWVEMWPLQEMPPELMKPDLFTYPCEMSLISQSTRL
                     NSYS"
BASE COUNT      340 a    229 c    226 g    396 t
ORIGIN

Alignment Scores:
Pred. No.:                7.09e-32       Length:          1191
Score:                    376.00         Matches:         69
Percent Similarity:       83.18%         Conservative:    20
Best Local Similarity:    64.49%         Mismatches:      18
Query Match:              68.49%         Indels:          0
DB:                       6              Gaps:            0

US-09-930-312-2 (1-107) x AX128507 (1-1191)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           |||:::|||||   |||||||:::|||||||||||||||||   :::|||||   |||
Db     673 GGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGT 732

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           :::   ::::::|||:::      ||||||:::||||||       :::|||:::
Db     733 GTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATC 792

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           :::|||   |||||||||||||||||||:::||||||:::|||||||||||||:::|||
Db     793 CTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTT 852

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           |||||||||   |||||||||   :::|||||||||||   |||:::|||||||||:::
Db     853 GTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTA 912

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           |||||||||   ||||||:::||||||||||||||||||||||||||||||||||||||
Db     913 GTGATTTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACA 972
```

```
Qy    101 AsnPhePheLysAspLysLeu 107
              ||||||:::   :::
Db    973 AGACCATTTAAAGAAATGATT 993

RESULT 13
AF190500
LOCUS       AF190500               2274 bp    mRNA    linear   PRI 10-AUG-2001
DEFINITION  Homo sapiens leucine-rich repeat-containing G protein-coupled
            receptor 7 (LGR7) mRNA, complete cds.
ACCESSION   AF190500
VERSION     AF190500.1  GI:10441729
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2274)
  AUTHORS   Hsu,S.Y., Kudo,M., Chen,T., Nakabayashi,K., Bhalla,A., van der
            Spek,P.J., van Duin,M. and Hsueh,A.J.
  TITLE     The three subfamilies of leucine-rich repeat-containing G
            protein-coupled receptors (LGR): identification of LGR6 and LGR7
            and the signaling mechanism for LGR7
  JOURNAL   Mol. Endocrinol. 14 (8), 1257-1271 (2000)
  MEDLINE   20388592
   PUBMED   10935549
REFERENCE   2  (bases 1 to 2274)
  AUTHORS   Hsu,S.Y. and Hsueh,A.J.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-SEP-1999) GYN/OB, Stanford University, MSOB S385,
            Stanford, CA 94305, USA
FEATURES             Location/Qualifiers
     source          1. .2274
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            1. .2274
                     /gene="LGR7"
     CDS             1. .2274
                     /gene="LGR7"
                     /codon_start=1
                     /product="leucine-rich repeat-containing G protein-coupled
                     receptor 7"
                     /protein_id="AAG17167.1"
                     /db_xref="GI:10441730"
                     /translation="MTSGSVFFYILIFGKYFSHGGGQDVKCSLGYFPCGNITKCLPQL
                     LHCNGVDDCGNQADEDNCGDNNGWSMQFDKYFASYYKMTSQYPFEAETPECLVGSVPV
                     QCLCQGLELDCDETNLRAVPSVSSNVTAMSLQWNLIRKLPPDCFKNYHDLQKLYLQNN
                     KITSISIYAFRGLNSLTKLYLSHNRITFLKPGVFEDLHRLEWLIIEDNHLSRISPPTF
                     YGLNSLILLVLMNNVLTRLPDKPLCQHMPRLHWLDLEGNHIHNLRNLTFISCSNLTVL
                     VMRKNKINHLNENTFAPLQKLDELDLGSNKIENLPPLIFKDLKELSQLNLSYNPIQKI
                     QANQFDYLVKLKSLSLEGIEISNIQQRMFRPLMNLSHIYFKKFQYCGYAPHVRSCKPN
                     TDGISSLENLLASIIQRVFVWVVSAVTCFGNIFVICMRPYIRSENKLYAMSIISLCCA
                     DCLMGIYLFVIGGFDLKFRGEYNKHAQLWMESTHCQLVGSLAILSTEVSVLLLTFLTL
                     EKYICIVYPFRCVRPGKCRTITVLILIWITGFIVAFIPLSNKEFFKNYYGTNGVCFPL
                     HSEDTESIGAQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEM
                     ILAKRFFFIVFTDALCWIPIFVVKFLSLLQVEIPGTITSWVVIFILPINSALNPILYT
                     LTTRPFKEMIHRFWYNYRQRKSMDSKGQKTYAPSFIWVEMWPLQEMPPELMKPDLFTY
                     PCEMSLISQSTRLNSYS"
BASE COUNT      669 a    460 c    421 g    724 t
ORIGIN

Alignment Scores:
Pred. No.:              1.3e-31       Length:        2274
Score:                  376.00        Matches:       69
Percent Similarity:     83.18%        Conservative:  20
Best Local Similarity:  64.49%        Mismatches:    18
Query Match:            68.49%        Indels:        0
DB:                     9             Gaps:          0

US-09-930-312-2 (1-107) x AF190500 (1-2274)

Qy      1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
          |||:::||||||   ||||||:::||||||||||||||||||||   :::||||||   |||
Db   1756 GGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGT 1815

Qy     21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
          :::   ::::::|||:::      ||||||:::||||||         :::|||:::
Db   1816 GTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATC 1875

Qy     41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
          :::|||   ||||||||||||||||||||||:::||||||:::||||||||||||:::|||
Db   1876 CTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTT 1935

Qy     61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
          |||||||||   |||||||||   :::||||||||||||   |||:::||||||||:::
Db   1936 GTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTA 1995

Qy     81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
          |||||||||   ||||||:::||||||||||||||||||||||||||||||||||||||||
Db   1996 GTGATTTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACA 2055

Qy    101 AsnPhePheLysAspLysLeu 107
            ||||||:::   :::
Db   2056 AGACCATTTAAAGAAATGATT 2076

RESULT 14
AX147820/c
LOCUS       AX147820               1018 bp    DNA     linear   PAT 08-JUN-2001
DEFINITION  Sequence 65 from Patent WO0136473.
ACCESSION   AX147820
VERSION     AX147820.1  GI:14346831
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1018)
  AUTHORS   Vogeli,G., Wood,L.S., Parodi,L.A., Hiebsch,R.R., Lind,P.,
            Slightom,J., Schellin,K.A., Kaytes,P.S., Bannigan,C.M., Ruff,V.,
            Sejlitz,T. and Huff,R.M.
  TITLE     Novel g protein-coupled receptors
  JOURNAL   Patent: WO 0136473-A 65 25-MAY-2001;
            PHARMACIA & UPJOHN COMPANY (US)
FEATURES             Location/Qualifiers
     source          1. .1018
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      332 a    179 c    189 g    318 t
ORIGIN

Alignment Scores:
Pred. No.:              2.14e-31      Length:        1018
Score:                  371.00        Matches:       73
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            67.58%        Indels:        0
DB:                     6             Gaps:          0

US-09-930-312-2 (1-107) x AX147820 (1-1018)

Qy      1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    539 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 480

Qy     21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    479 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 420

Qy     41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    419 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 360

Qy     61 ValValLysIleLeuSerLeuPheArgValGluIlePro 73
          |||||||||||||||||||||||||||||||||||||
Db    359 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCA 321
```

RESULT 15
AL138708
LOCUS AL138708 170522 bp DNA linear PRI 01-FEB-2001
DEFINITION Human DNA sequence from clone RP11-432E15 on chromosome 13 Contains STSs, GSSs and a CpG island, complete sequence.
ACCESSION AL138708
VERSION AL138708.17 GI:11137618
KEYWORDS HTG; CpG island.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 170522)
AUTHORS Phillimore,B.
TITLE Direct Submission
JOURNAL Submitted (01-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT On Nov 10, 2000 this sequence version replaced gi:11121365.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 13, constructed by the Sanger Centre Chromosome 13 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr13
This sequence is the entire insert of clone RP11-432E15 The true left end of clone RP11-154I23 is at 137612 in this sequence. The true right end of clone RP11-95N14 is at 13246 in this sequence. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated repeat sequence elements. Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key. RP11-432E15 is from the library RPCI-11.2 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6.
FEATURES Location/Qualifiers
source 1. .170522
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
/clone="RP11-432E15"
/clone_lib="RPCI-11.2"
misc_feature 19. .289
/note="match: GSS: Em:AQ549338"
repeat_region 194. .299
/note="53 copies 2 mer tg 71% conserved"
repeat_region 208. .299
/note="23 copies 4 mer tgtg 77% conserved"
repeat_region 1456. .1750
/note="AluSx repeat: matches 3. .300 of consensus"
repeat_region 2078. .2183
/note="MIR repeat: matches 6. .114 of consensus"
misc_feature 2600. .3122
/note="match: GSS: Em:AQ529216"
repeat_region 2646. .2940
/note="AluSg repeat: matches 1. .302 of consensus"
repeat_region 3590. .3894
/note="AluSq repeat: matches 1. .296 of consensus"
repeat_region 3926. .4462
/note="L2 repeat: matches 2233. .2750 of consensus"
repeat_region 4658. .4805
/note="L2 repeat: matches 2355. .2503 of consensus"
repeat_region 4911. .5224
/note="AluSq repeat: matches 1. .312 of consensus"
repeat_region 5237. .5332
/note="L2 repeat: matches 2575. .2694 of consensus"
repeat_region 5788. .6085
/note="AluJo repeat: matches 3. .301 of consensus"
misc_feature 7355. .7747
/note="sequence from overlapping clone bA159O9 (AL136106).Assembly confirmed by restriction digest."
repeat_region 7388. .7487
/note="50 copies 2 mer at 63% conserved"
repeat_region 7396. .7487
/note="23 copies 4 mer atat 66% conserved"
repeat_region 7884. .7915
/note="MER20 repeat: matches 30. .62 of consensus"
repeat_region 7916. .8219
/note="AluJb repeat: matches 1. .302 of consensus"
repeat_region 8220. .8376
/note="MER20 repeat: matches 62. .210 of consensus"
misc_feature 8975. .8985
/note="sequence from overlapping clone bA159O9(AL136106).Assembly confirmed by restriction digest."
repeat_region 9078. .9121
/note="11 copies 4 mer caca 100% conserved"
misc_feature complement(9603. .10104)
/note="match: GSS: Em:AQ569313"
misc_feature 11234. .11331
/note="sequence from overlapping clone bA159O9 (AL136106).Assembly confirmed by restriction digest."
misc_feature complement(11953. .12400)
/note="match: GSS: Em:AQ486928"
repeat_region 12477. .12601
/note="MIR repeat: matches 15. .138 of consensus"
misc_feature 12713. .12980
/note="sequence from overlapping clone bA159O9 (AL136106).Assembly confirmed by restriction digest."
repeat_region 13272. .13417
/note="MIR repeat: matches 38. .188 of consensus"
misc_feature complement(13670. .13919)
/note="match: GSS: Em:AQ021664"
misc_feature complement(13740. .14130)
/note="match: GSS: Em:AQ032630"
misc_feature 13771. .14227
/note="sequence from overlapping clone bA159O9 (AL136106).Assembly confirmed by restriction digest."
repeat_region 13864. .14087
/note="L2 repeat: matches 2515. .2750 of consensus"
repeat_region 16061. .16342
/note="MER1B repeat: matches 1. .244 of consensus"
repeat_region 16343. .16655
/note="AluSx repeat: matches 1. .312 of consensus"
repeat_region 16656. .16680
/note="MER1B repeat: matches 244. .305 of consensus"
misc_feature complement(17019. .17275)
/note="match: GSS: Em:AQ452676"
misc_feature 17070. .17578
/note="match: GSS: Em:AQ511484"
misc_feature 17087. .17515
/note="match: GSS: Em:AQ817416"
misc_feature 17278. .17794
/note="match: GSS: Em:AQ358157"
repeat_region 18301. .18381
/note="L1MB2 repeat: matches 6082. .6166 of consensus"
misc_feature complement(18563. .19010)
/note="match: GSS: Em:AQ242897"
repeat_region 20403. .20559
/note="MER20 repeat: matches 1. .153 of consensus"
repeat_region 20870. .21069
/note="100 copies 2 mer tt 69% conserved"
repeat_region 20937. .21076
/note="35 copies 4 mer tttc 72% conserved"
repeat_region 21080. .21392

```
                /note="AluJb repeat: matches 1. .310 of consensus"
     repeat_region   21395. .21538
                /note="MIR repeat: matches 94. .245 of consensus"
     repeat_region   23716. .23960
                /note="L2 repeat: matches 2381. .2617 of consensus"
     repeat_region   23961. .24213
                /note="AluSx repeat: matches 47. .299 of consensus"
     repeat_region   24214. .24342
                /note="L2 repeat: matches 2617. .2747 of consensus"
     repeat_region   25636. .25785
                /note="MER5A repeat: matches 32. .189 of consensus"
     repeat_region   25638. .25810
                /note="MER5A repeat: matches 17. .189 of consensus"
     repeat_region   26013. .26227
                /note="MIR repeat: matches 10. .256 of consensus"
     repeat_region   28426. .28484
                /note="L2 repeat: matches 2641. .2700 of consensus"
     repeat_region   28639. .28960
                /note="AluSq repeat: matches 1. .310 of consensus"
     repeat_region   29689. .29861
                /note="L2 repeat: matches 2006. .2186 of consensus"
     misc_feature    complement(29720. .30187)
                /note="match: GSS: Em:AQ317477"
     repeat_region   29901. .30075
                /note="MER5A repeat: matches 1. .184 of consensus"
     misc_feature    complement(30119. .30414)
                /note="match: GSS: Em:AQ230143"
     repeat_region   32462. .32774
                /note="AluSx repeat: matches 1. .312 of consensus"
     repeat_region   32807. .32877
                /note="L2 repeat: matches 2635. .2707 of consensus"
     repeat_region   35365. .35806
                /note="L2 repeat: matches 2248. .2742 of consensus"
     misc_feature    complement(35852. .36413)
                /note="match: GSS: Em:AQ284653
                match: STS: Em:G55301"
     misc_feature    complement(36043. .36433)
                /note="match: GSS: Em:AQ816559"
     repeat_region   36420. .36469
                /note="MIR repeat: matches 96. .148 of consensus"
     misc_feature    36429. .36961
                /note="match: GSS: Em:AQ394779"
     misc_feature    36452. .36845
                /note="match: GSS: Em:AQ212353"
     repeat_region   39203. .39242
                /note="20 copies 2 mer ca 100% conserved"
     repeat_region   39250. .39648
                /note="L2 repeat: matches 2307. .2710 of consensus"
     repeat_region   39703. .39750
                /note="24 copies 2 mer ca 75% conserved"
     repeat_region   39859. .39976
                /note="L1MC3 repeat: matches 7621. .7739 of consensus"
     repeat_region   39988. .40283
                /note="AluSx repeat: matches 1. .296 of consensus"
     repeat_region   40873. .41144
                /note="AluY repeat: matches 39. .310 of consensus"

Alignment Scores:
Pred. No.:              2.68e-29        Length:         170522
Score:                  371.00          Matches:        73
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            67.58%          Indels:         0
DB:                     9               Gaps:           0

US-09-930-312-2 (1-107) x AL138708 (1-170522)

Qy      1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  42034 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 42093

Qy     21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db  42094 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 42153

Qy     41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  42154 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 42213

Qy     61 ValValLysIleLeuSerLeuPheArgValGluIlePro 73
          |||||||||||||||||||||||||||||||||||||||
Db  42214 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCA 42252


Search completed: May 26, 2003, 20:51:48
Job time : 2401 secs
```

GenCore version 5.1.4_p5_4578

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 26, 2003, 19:02:42 ; Search time 236 Seconds
(without alignments)
1021.033 Million cell updates/sec

Title: US-09-930-312-2
Perfect score: 549
Sequence: 1 GVNLLAFLIIVFSYITMFCS.....NSALNPILYTLTTNFFKDKL 107

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09930312/runat_22052003_114612_27057/app_query.fasta_1.263
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09930312_@CGN_1_1_263_@runat_22052003_114612_27057 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	549	100.0	321	24	AAD32027	Human novel G-prot
	2	549	100.0	1065	24	ABL40194	Human G protein-co
	3	549	100.0	1068	22	AAS07943	Human cDNA encodin
	4	549	100.0	1473	24	ABL40192	Human G protein-co
	5	549	100.0	1545	24	ABL40191	Human G protein-co
	6	549	100.0	1830	24	ABL40188	Human G protein-co
	7	549	100.0	2142	24	ABK51947	cDNA encoding huma
	8	549	100.0	2190	24	ABL40197	Human G protein-co
	9	549	100.0	2214	24	ABK51944	cDNA encoding huma
	10	549	100.0	2262	24	ABL40196	Human G protein-co
	11	376	68.5	530	22	AAI99584	Human expressed po
	12	376	68.5	530	22	AAL35644	Human musculoskele
	13	376	68.5	530	22	ABA06471	Human cDNA SEQ ID
	14	376	68.5	530	22	AAS28950	cDNA encoding for
	15	376	68.5	530	22	AAS29573	Human endocrine po
	16	376	68.5	530	22	AAS30187	DNA encoding rena
	17	376	68.5	530	22	AAS34845	cDNA encoding nove
	18	376	68.5	530	23	ABK43875	DNA encoding novel
	19	376	68.5	1089	22	AAI99557	Human expressed po
c	20	376	68.5	1162	23	ABK43573	DNA encoding novel
	21	376	68.5	1191	22	AAD06507	Human CON222 G pro
	22	376	68.5	1804	21	AAZ90524	Human GPCR protein
	23	376	68.5	2467	20	AAZ25345	Human LGR7 long fo
	24	376	68.5	3584	20	AAZ25346	Human LGR7 short f
c	25	371	67.6	1018	22	AAH51001	Human nGPCR57 codi
	26	370	67.4	420	21	AAC75735	Human ORFX ORF1290
	27	349	63.6	474	22	AAF28059	Human HGRL101 G-pr
	28	234	42.6	1015	22	AAS57085	cDNA encoding Dros
	29	234	42.6	1015	23	ABL07293	Drosophila melanog
	30	225	41.0	1080	22	AAS57089	cDNA encoding Dros
	31	225	41.0	1080	23	ABL08799	Drosophila melanog
	32	225	41.0	3187	22	AAS57088	DNA encoding Droso
	33	225	41.0	3187	23	ABL08798	Drosophila melanog
	34	225	41.0	3494	23	ABL08218	Drosophila melanog
	35	181	33.0	3574	22	AAS57084	DNA encoding Droso
	36	181	33.0	3574	23	ABL07292	Drosophila melanog
	37	158.5	28.9	3578	23	ABL03716	Drosophila melanog
	38	142.5	26.0	2115	12	AAQ11120	Exon 4 of Human lu
	39	142	25.9	2586	22	AAH49414	D. melanogaster pe
	40	142	25.9	3909	23	ABL14533	Drosophila melanog
c	41	142	25.9	12248	23	ABL14532	Drosophila melanog
	42	138.5	25.2	1831	16	AAQ97312	Rat A3 adenosine r
	43	136.5	24.9	2902	11	AAQ06633	Clone encoding rat
	44	135.5	24.7	2987	13	AAQ22937	Sequence encoding
	45	128	23.3	1681	22	AAH17465	Human cDNA sequenc

ALIGNMENTS

RESULT 1
AAD32027
ID AAD32027 standard; DNA; 321 BP.
XX
AC AAD32027;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human novel G-protein coupled receptor (nGPCR)-1079 DNA.
XX
KW Human; novel G-protein coupled receptor; nGPCR-1079; viral infection;
KW gene; therapy; human immunodeficiency virus; HIV; pain; migraine;
KW central nervous system disorder; stroke; manic depression; obesity;
KW metabolic disorder; anorexia; cardiovascular disorder; type 2 diabetes;
KW cancer; myocardial infarction; hypotension; degenerative disorder;
KW Parkinson's disease; Alzheimer's disease; neurological disorder;
KW schizophrenia; anxiety; inflammatory condition; rheumatoid arthritis;
KW thyroid disorder; autoimmune disorder; hormonal disorder; renal failure;

```
KW   psoriasis; movement disorder; analgesic; cytostatic; neuroprotective;
KW   anorectic; hypotensive; hypertensive; tranquilliser; anticonvulsant;
KW   metabolic; neuroleptic; thrombolytic; cardiant; immunosuppressive; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..321
FT                   /*tag= a
FT                   /product= "Human nGPCR-1079 protein"
FT                   /note= "CDS does not include start and stop codon"
FT                   /partial
XX
PN   WO200214496-A2.
XX
PD   21-FEB-2002.
XX
PF   15-AUG-2001; 2001WO-US25513.
XX
PR   15-AUG-2000; 2000US-225262P.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Lind P;
XX
DR   WPI; 2002-269192/31.
DR   P-PSDB; AAE20148.
XX
PT   An isolated nucleic acid molecule encoding novel G-protein coupled
PT   receptor polypeptide which is useful for treating obesity, diabetes,
PT   Parkinson's disease, manic depression, migraine, rheumatoid arthritis
PT   -
XX
PS   Claim 4; Page 60; 93pp; English.
XX
CC   The patent discloses novel G-protein coupled receptor (nGPCR)-1079 genes
CC   and their corresponding proteins. Sequences of the invention are useful
CC   for treating diseases such as viral infections caused by human immuno-
CC   deficiency virus (HIV)-1 or 2, central nervous system (CNS) disorders
CC   (e.g. pain, including migraine, stroke, manic depression), metabolic
CC   disorders (e.g. obesity and anorexia), cancers, cardiovascular disorders
CC   (e.g. type 2 diabetes, myocardial infarction, hypotension), degenerative
CC   disorders (e.g. Parkinson's disease, Alzheimer's disease), neurological
CC   disorders (e.g. schizophrenia and anxiety), inflammatory conditions,
CC   rheumatoid arthritis, thyroid disorders, autoimmune disorders, hormonal
CC   disorders, renal failure, psoriasis and movement disorders. The present
CC   sequence is a DNA encoding human nGPCR-1079 protein.
XX
SQ   Sequence 321 BP; 76 A; 67 C; 59 G; 119 T; 0 other;

Alignment Scores:
Pred. No.:              3.11e-58      Length:          321
Score:                  549.00        Matches:         107
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            100.00%       Indels:          0
DB:                     24            Gaps:            0

US-09-930-312-2 (1-107) x AAD32027 (1-321)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300

Qy     101 AsnPhePheLysAspLysLeu 107
           |||||||||||||||||||||
Db     301 AACTTTTTTAAGGACAAGTTG 321

RESULT 2
ABL40194
ID   ABL40194 standard; cDNA; 1065 BP.
XX
AC   ABL40194;
XX
DT   23-MAY-2002  (first entry)
XX
DE   Human G protein-coupled receptor TGR17-4 encoding cDNA SEQ ID NO:11.
XX
KW   Human; G protein-coupled receptor; TGR17-4; nootropic; antiinflammatory;
KW   vasotropic; immunomodulator; cytostatic; gene therapy; protein therapy;
KW   neurological; inflammatory; circulatory; degenerative; immune system;
KW   digestive disease; cancer; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1065
FT                   /*tag= a
FT                   /partial
FT                   /product= "TGR17-4"
FT                   /note= "no stop codon given"
XX
PN   WO200204640-A1.
XX
PD   17-JAN-2002.
XX
PF   06-JUL-2001; 2001WO-JP05878.
XX
PR   07-JUL-2000; 2000JP-0211989.
PR   18-DEC-2000; 2000JP-0383794.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
PI   Moriya T,  Ito T,  Shintani Y,  Miyajima N;
XX
DR   WPI; 2002-179706/23.
DR   P-PSDB; ABB06253.
XX
PT   G-protein coupled receptor protein TGR17 of human origin and DNA
PT   encoding it for diagnosis and treatment of cancer and circulatory and
PT   other diseases associated with its expression -
XX
PS   Claim 7; Page 125-126; 145pp; Japanese.
XX
CC   The present invention describes a human guanine nucleotide binding
CC   protein (G protein) coupled receptor protein designated TGR17, which has
CC   5 variant forms (TGR17-1 to TGR17-5) of the receptor protein having
CC   additional residues at the N-terminal. The TGR17 proteins have nootropic,
CC   antiinflammatory, vasotropic, immunomodulator and cytostatic activities.
CC   The TGR17 polynucleotides and protein can be used in gene therapy and
CC   protein therapy. G-protein coupled receptor proteins are cell membrane
CC   proteins mediating the cellular response to a large variety of signalling
CC   molecules. The TGR17 polynucleotides and proteins can be used in the
CC   diagnosis, treatment and prevention of diseases including neurological,
CC   inflammatory, circulatory, degenerative, immune system and digestive
CC   diseases and cancer. The present sequence encodes human TGR17-4 from the
CC   present invention.
XX
SQ   Sequence 1065 BP; 277 A; 214 C; 213 G; 361 T; 0 other;
```

```
Alignment Scores:
Pred. No.:              1.47e-57        Length:         1065
Score:                  549.00          Matches:        107
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     24              Gaps:           0

US-09-930-312-2 (1-107) x ABL40194 (1-1065)

Qy      1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    589 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 648

Qy     21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 708

Qy     41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 768

Qy     61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 828

Qy     81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    829 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 888

Qy    101 AsnPhePheLysAspLysLeu 107
          |||||||||||||||||||||
Db    889 AACTTTTTTAAGGACAAGTTG 909

RESULT 3
AAS07943
ID   AAS07943 standard; cDNA; 1068 BP.
XX
AC   AAS07943;
XX
DT   23-OCT-2001  (first entry)
XX
DE   Human cDNA encoding G-protein coupled receptor, hRUP16.
XX
KW   Human; G-protein coupled receptor; GPCR; hRUP16; agonist;
KW   inverse agonist; lung cancer; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1068
FT                   /*tag= a
FT                   /product= "hRUP16"
XX
PN   WO200136471-A2.
XX
PD   25-MAY-2001.
XX
PF   16-NOV-2000; 2000WO-US31509.
XX
PR   17-NOV-1999;   99US-0166088.
PR   17-NOV-1999;   99US-0166099.
PR   17-NOV-1999;   99US-0166369.
PR   23-DEC-1999;   99US-0171900.
PR   23-DEC-1999;   99US-0171901.
PR   23-DEC-1999;   99US-0171902.
PR   11-FEB-2000; 2000US-0181749.
PR   14-MAR-2000; 2000US-0189258.
PR   14-MAR-2000; 2000US-0189259.
PR   10-APR-2000; 2000US-0195898.
PR   10-APR-2000; 2000US-0195899.
PR   10-APR-2000; 2000US-0196078.
PR   28-APR-2000; 2000US-0200419.
PR   12-MAY-2000; 2000US-0203630.
PR   12-JUN-2000; 2000US-0210741.
PR   12-JUN-2000; 2000US-0210982.
PR   21-AUG-2000; 2000US-0226760.
PR   26-SEP-2000; 2000US-0235418.
PR   26-SEP-2000; 2000US-0235779.
PR   20-OCT-2000; 2000US-0242332.
PR   20-OCT-2000; 2000US-0242343.
XX
PA   (AREN-) ARENA PHARM INC.
XX
PI   Chen R,  Dang HT,  Lowitz KP;
XX
DR   WPI; 2001-355616/37.
DR   P-PSDB; AAU04370.
XX
PT   Endogenous and non-endogenous versions of human G-protein coupled
PT   receptors for direct identification of candidate compounds as agonists,
PT   inverse agonists or partial agonists for use as therapeutic agents -
XX
PS   Claim 35; Page 104-105; 159pp; English.
XX
CC   The sequence encodes a human G-protein coupled receptor (GPCR),
CC   hRUP16 The endogenous and non-endogenous, constitutively activated
CC   versions of human G-protein coupled receptors (GPCR), are useful for
CC   direct identification of candidate compounds as receptor agonists,
CC   inverse agonists or partial agonists having applicability as therapeutic
CC   agents for treating diseases related to GPCR, e.g. lung cancer.
CC   Non-endogenous version of human GPCRs are also utilized in research
CC   settings and in vitro and in vivo system, incorporating GPCRs can be
CC   utilised to elucidate and understand the roles these receptors
CC   play in the human condition, both normal and diseased.
XX
SQ   Sequence 1068 BP; 278 A; 215 C; 214 G; 361 T; 0 other;

Alignment Scores:
Pred. No.:              1.47e-57        Length:         1068
Score:                  549.00          Matches:        107
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     22              Gaps:           0

US-09-930-312-2 (1-107) x AAS07943 (1-1068)

Qy      1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    589 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 648

Qy     21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    649 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 708

Qy     41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    709 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 768

Qy     61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    769 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 828

Qy     81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    829 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 888

Qy    101 AsnPhePheLysAspLysLeu 107
          |||||||||||||||||||||
Db    889 AACTTTTTTAAGGACAAGTTG 909

RESULT 4
ABL40192
ID   ABL40192 standard; cDNA; 1473 BP.
XX
```

```
AC   ABL40192;
XX
DT   23-MAY-2002  (first entry)
XX
DE   Human G protein-coupled receptor TGR17-3 encoding cDNA SEQ ID NO:8.
XX
KW   Human; G protein-coupled receptor; TGR17-3; nootropic; antiinflammatory;
KW   vasotropic; immunomodulator; cytostatic; gene therapy; protein therapy;
KW   neurological; inflammatory; circulatory; degenerative; immune system;
KW   digestive disease; cancer; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1473
FT                   /*tag= a
FT                   /partial
FT                   /product= "TGR17-3"
FT                   /note= "no stop codon given"
XX
PN   WO200204640-A1.
XX
PD   17-JAN-2002.
XX
PF   06-JUL-2001; 2001WO-JP05878.
XX
PR   07-JUL-2000; 2000JP-0211989.
PR   18-DEC-2000; 2000JP-0383794.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
PI   Moriya T,  Ito T,  Shintani Y,  Miyajima N;
XX
DR   WPI; 2002-179706/23.
DR   P-PSDB; ABB06252.
XX
PT   G-protein coupled receptor protein TGR17 of human origin and DNA
PT   encoding it for diagnosis and treatment of cancer and circulatory and
PT   other diseases associated with its expression -
XX
PS   Claim 7; Page 122-123; 145pp; Japanese.
XX
CC   The present invention describes a human guanine nucleotide binding
CC   protein (G protein) coupled receptor protein designated TGR17, which has
CC   5 variant forms (TGR17-1 to TGR17-5) of the receptor protein having
CC   additional residues at the N-terminal. The TGR17 proteins have nootropic,
CC   antiinflammatory, vasotropic, immunomodulator and cytostatic activities.
CC   The TGR17 polynucleotides and protein can be used in gene therapy and
CC   protein therapy. G-protein coupled receptor proteins are cell membrane
CC   proteins mediating the cellular response to a large variety of signalling
CC   molecules. The TGR17 polynucleotides and proteins can be used in the
CC   diagnosis, treatment and prevention of diseases including neurological,
CC   inflammatory, circulatory, degenerative, immune system and digestive
CC   diseases and cancer. The present sequence encodes human TGR17-3 from the
CC   present invention.
XX
SQ   Sequence 1473 BP; 405 A; 310 C; 278 G; 480 T; 0 other;

Alignment Scores:
Pred. No.:              2.23e-57      Length:          1473
Score:                  549.00        Matches:         107
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            100.00%       Indels:          0
DB:                     24            Gaps:            0

US-09-930-312-2 (1-107) x ABL40192 (1-1473)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     997 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1056

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1057 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1116

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1117 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1176

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1177 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1236

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1237 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 1296

Qy     101 AsnPhePheLysAspLysLeu 107
           |||||||||||||||||||||
Db    1297 AACTTTTTTAAGGACAAGTTG 1317

RESULT 5
ABL40191
ID   ABL40191 standard; cDNA; 1545 BP.
XX
AC   ABL40191;
XX
DT   23-MAY-2002  (first entry)
XX
DE   Human G protein-coupled receptor TGR17-2 encoding cDNA SEQ ID NO:6.
XX
KW   Human; G protein-coupled receptor; TGR17-2; nootropic; antiinflammatory;
KW   vasotropic; immunomodulator; cytostatic; gene therapy; protein therapy;
KW   neurological; inflammatory; circulatory; degenerative; immune system;
KW   digestive disease; cancer; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1545
FT                   /*tag= a
FT                   /partial
FT                   /product= "TGR17-2"
FT                   /note= "no stop codon given"
XX
PN   WO200204640-A1.
XX
PD   17-JAN-2002.
XX
PF   06-JUL-2001; 2001WO-JP05878.
XX
PR   07-JUL-2000; 2000JP-0211989.
PR   18-DEC-2000; 2000JP-0383794.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
PI   Moriya T,  Ito T,  Shintani Y,  Miyajima N;
XX
DR   WPI; 2002-179706/23.
DR   P-PSDB; ABB06251.
XX
PT   G-protein coupled receptor protein TGR17 of human origin and DNA
PT   encoding it for diagnosis and treatment of cancer and circulatory and
PT   other diseases associated with its expression -
XX
PS   Claim 7; Page 118-119; 145pp; Japanese.
XX
CC   The present invention describes a human guanine nucleotide binding
CC   protein (G protein) coupled receptor protein designated TGR17, which has
CC   5 variant forms (TGR17-1 to TGR17-5) of the receptor protein having
CC   additional residues at the N-terminal. The TGR17 proteins have nootropic,
CC   antiinflammatory, vasotropic, immunomodulator and cytostatic activities.
CC   The TGR17 polynucleotides and protein can be used in gene therapy and
CC   protein therapy. G-protein coupled receptor proteins are cell membrane
CC   proteins mediating the cellular response to a large variety of signalling
```

```
CC   molecules. The TGR17 polynucleotides and proteins can be used in the
CC   diagnosis, treatment and prevention of diseases including neurological,
CC   inflammatory, circulatory, degenerative, immune system and digestive
CC   diseases and cancer. The present sequence encodes human TGR17-2 from the
CC   present invention.
XX
SQ   Sequence 1545 BP; 429 A; 320 C; 290 G; 506 T; 0 other;

Alignment Scores:
Pred. No.:              2.37e-57      Length:          1545
Score:                  549.00        Matches:         107
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            100.00%       Indels:          0
DB:                     24            Gaps:            0

US-09-930-312-2 (1-107) x ABL40191 (1-1545)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1069 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1128

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1129 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1188

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1189 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1248

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1249 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1308

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1309 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 1368

Qy     101 AsnPhePheLysAspLysLeu 107
           |||||||||||||||||||||
Db    1369 AACTTTTTTAAGGACAAGTTG 1389

RESULT 6
ABL40188
ID   ABL40188 standard; cDNA; 1830 BP.
XX
AC   ABL40188;
XX
DT   23-MAY-2002  (first entry)
XX
DE   Human G protein-coupled receptor TGR17-1 encoding cDNA SEQ ID NO:2.
XX
KW   Human; G protein-coupled receptor; TGR17-1; nootropic; antiinflammatory;
KW   vasotropic; immunomodulator; cytostatic; gene therapy; protein therapy;
KW   neurological; inflammatory; circulatory; degenerative; immune system;
KW   digestive disease; cancer; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1830
FT                   /*tag=  a
FT                   /partial
FT                   /product= "TGR17-1"
FT                   /note= "no stop codon given"
XX
PN   WO200204640-A1.
XX
PD   17-JAN-2002.
XX
PF   06-JUL-2001; 2001WO-JP05878.
XX
PR   07-JUL-2000; 2000JP-0211989.
PR   18-DEC-2000; 2000JP-0383794.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
PI   Moriya T,  Ito T,  Shintani Y,  Miyajima N;
XX
DR   WPI; 2002-179706/23.
DR   P-PSDB; ABB06250.
XX
PT   G-protein coupled receptor protein TGR17 of human origin and DNA
PT   encoding it for diagnosis and treatment of cancer and circulatory and
PT   other diseases associated with its expression -
XX
PS   Claim 7; Page 114-115; 145pp; Japanese.
XX
CC   The present invention describes a human guanine nucleotide binding
CC   protein (G protein) coupled receptor protein designated TGR17, which has
CC   5 variant forms (TGR17-1 to TGR17-5) of the receptor protein having
CC   additional residues at the N-terminal. The TGR17 proteins have nootropic,
CC   antiinflammatory, vasotropic, immunomodulator and cytostatic activities.
CC   The TGR17 polynucleotides and protein can be used in gene therapy and
CC   protein therapy. G-protein coupled receptor proteins are cell membrane
CC   proteins mediating the cellular response to a large variety of signalling
CC   molecules. The TGR17 polynucleotides and proteins can be used in the
CC   diagnosis, treatment and prevention of diseases including neurological,
CC   inflammatory, circulatory, degenerative, immune system and digestive
CC   diseases and cancer. The present sequence encodes human TGR17-1 from the
CC   present invention.
XX
SQ   Sequence 1830 BP; 516 A; 377 C; 351 G; 586 T; 0 other;

Alignment Scores:
Pred. No.:              2.95e-57      Length:          1830
Score:                  549.00        Matches:         107
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            100.00%       Indels:          0
DB:                     24            Gaps:            0

US-09-930-312-2 (1-107) x ABL40188 (1-1830)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1354 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1413

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1414 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1473

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1474 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1533

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1534 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1593

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1594 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 1653

Qy     101 AsnPhePheLysAspLysLeu 107
           |||||||||||||||||||||
Db    1654 AACTTTTTTAAGGACAAGTTG 1674

RESULT 7
ABK51947
ID   ABK51947 standard; cDNA; 2142 BP.
XX
AC   ABK51947;
XX
DT   27-AUG-2002  (first entry)
XX
```

```
DE   cDNA encoding human G-protein coupled receptor HGPRBMY5 splice variant.
XX
KW   Human; G-protein coupled receptor; GPCR; HGPRBMY5; colon; brain;
KW   ovary; thymus; lung; immune system; cancer; immune disorder;
KW   neurological disorder; infection; human immunodeficiency virus; HIV;
KW   antiallergic; antiasthmatic; dermatological; antiarteriosclerotic;
KW   antiulcer; antidiabetic; nephrotropic; osteopathic; antiarthritic;
KW   antiinflammatory; antirheumatic; antithyroid; cytostatic; vulnerary;
KW   virucide; antibacterial; antifungal; antiparasitic; protozoacide;
KW   antihelminthic; nootropic; neuroprotective; antidepressant;
KW   anticonvulsant; antiparkinsonian; neuroleptic; anti-HIV; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..2142
FT                   /*tag= a
FT                   /product= "GPCR HGPRBMY5"
XX
PN   WO200226824-A2.
XX
PD   04-APR-2002.
XX
PF   26-SEP-2001; 2001WO-US30365.
XX
PR   27-SEP-2000; 2000US-235713P.
PR   16-JAN-2001; 2001US-261781P.
PR   19-JUL-2001; 2001US-306605P.
PR   03-AUG-2001; 2001US-310436P.
XX
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI   Feder JN,  Mintier G,  Ramanathan CS,  Hawken DR;
XX
DR   WPI; 2002-435196/46.
DR   P-PSDB; AAU97159.
XX
PT   Novel G protein-coupled receptor, HGPRBMY5 polypeptide, useful for
PT   treatment of AIDS, allergies, asthma, atherosclerosis, ulcerative
PT   colitis, atopic dermatitis, diabetes mellitus, glomerulonephritis,
PT   osteoarthritis
XX
PS   Claim 1; Fig 5; 148pp; English.
XX
CC   The present invention relates to the isolation of a novel human
CC   G-protein coupled receptor (GPCR) (HGPRBMY5), and the polynucleotide
CC   sequence encoding it. The HGPRBMY5 polypeptide and polynucleotide
CC   are useful for preventing, treating or ameliorating a disease,
CC   disorder or condition related to the colon, brain, ovaries, thymus,
CC   lungs or immune system. They are particularly useful for the
CC   treatment or prevention of cancers, immune disorders, neurological
CC   disorders, and diseases related to the brain, ovaries, thymus or
CC   lungs. The polynucleotide sequence is useful for diagnosing or
CC   determining susceptibility to infections such as bacterial, fungal,
CC   protozoan and viral infections, particularly infections caused by
CC   human immunodeficiency virus (HIV or HIV-2). The present sequence
CC   encodes a splice variant of human GPCR HGPRBMY5.
XX
SQ   Sequence 2142 BP; 613 A; 435 C; 408 G; 686 T; 0 other;

Alignment Scores:
Pred. No.:              3.61e-57      Length:          2142
Score:                  549.00        Matches:         107
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            100.00%       Indels:          0
DB:                     24            Gaps:            0

US-09-930-312-2 (1-107) x ABK51947 (1-2142)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1663 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1722

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1723 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1782

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1783 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1842

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1843 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1902

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1903 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 1962

Qy     101 AsnPhePheLysAspLysLeu 107
           |||||||||||||||||||||
Db    1963 AACTTTTTTAAGGACAAGTTG 1983

RESULT 8
ABL40197
ID   ABL40197 standard; cDNA; 2190 BP.
XX
AC   ABL40197;
XX
DT   23-MAY-2002  (first entry)
XX
DE   Human G protein-coupled receptor TGR17-6 encoding cDNA SEQ ID NO:16.
XX
KW   Human; G protein-coupled receptor; TGR17-6; nootropic; antiinflammatory;
KW   vasotropic; immunomodulator; cytostatic; gene therapy; protein therapy;
KW   neurological; inflammatory; circulatory; degenerative; immune system;
KW   digestive disease; cancer; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..2190
FT                   /*tag= a
FT                   /partial
FT                   /product= "TGR17-6"
FT                   /note= "no stop codon given"
XX
PN   WO200204640-A1.
XX
PD   17-JAN-2002.
XX
PF   06-JUL-2001; 2001WO-JP05878.
XX
PR   07-JUL-2000; 2000JP-0211989.
PR   18-DEC-2000; 2000JP-0383794.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
PI   Moriya T,  Ito T,  Shintani Y,  Miyajima N;
XX
DR   WPI; 2002-179706/23.
DR   P-PSDB; ABB06255.
XX
PT   G-protein coupled receptor protein TGR17 of human origin and DNA
PT   encoding it for diagnosis and treatment of cancer and circulatory and
PT   other diseases associated with its expression -
XX
PS   Claim 7; Page 135-136; 145pp; Japanese.
XX
CC   The present invention describes a human guanine nucleotide binding
CC   protein (G protein) coupled receptor protein designated TGR17, which has
CC   5 variant forms (TGR17-1 to TGR17-5) of the receptor protein having
CC   additional residues at the N-terminal. The TGR17 proteins have nootropic,
CC   antiinflammatory, vasotropic, immunomodulator and cytostatic activities.
CC   The TGR17 polynucleotides and protein can be used in gene therapy and
```

```
CC   protein therapy. G-protein coupled receptor proteins are cell membrane
CC   proteins mediating the cellular response to a large variety of signalling
CC   molecules. The TGR17 polynucleotides and proteins can be used in the
CC   diagnosis, treatment and prevention of diseases including neurological,
CC   inflammatory, circulatory, degenerative, immune system and digestive
CC   diseases and cancer. The present sequence encodes human TGR17-6 from the
CC   present invention.
XX
SQ   Sequence 2190 BP; 626 A; 457 C; 409 G; 698 T; 0 other;

Alignment Scores:
Pred. No.:              3.72e-57      Length:         2190
Score:                  549.00        Matches:        107
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            100.00%       Indels:         0
DB:                     24            Gaps:           0

US-09-930-312-2 (1-107) x ABL40197 (1-2190)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1714 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1773

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1774 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1833

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1834 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1893

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1894 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1953

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1954 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2013

Qy     101 AsnPhePheLysAspLysLeu 107
           |||||||||||||||||||||
Db    2014 AACTTTTTTAAGGACAAGTTG 2034

RESULT 9
ABK51944
ID   ABK51944 standard; cDNA; 2214 BP.
XX
AC   ABK51944;
XX
DT   27-AUG-2002  (first entry)
XX
DE   cDNA encoding human G-protein coupled receptor HGPRBMY5.
XX
KW   Human; G-protein coupled receptor; GPCR; HGPRBMY5; colon; brain;
KW   ovary; thymus; lung; immune system; cancer; immune disorder;
KW   neurological disorder; infection; human immunodeficiency virus; HIV;
KW   antiallergic; antiasthmatic; dermatological; antiarteriosclerotic;
KW   antiulcer; antidiabetic; nephrotropic; osteopathic; antiarthritic;
KW   antiinflammatory; antirheumatic; antithyroid; cytostatic; vulnerary;
KW   virucide; antibacterial; antifungal; antiparasitic; protozoacide;
KW   antihelminthic; nootropic; neuroprotective; antidepressant;
KW   anticonvulsant; antiparkinsonian; neuroleptic; anti-HIV; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..2214
FT                   /*tag= a
FT                   /product= "GPCR HGPRBMY5"
XX
PN   WO200226824-A2.
XX
PD   04-APR-2002.
XX
PF   26-SEP-2001; 2001WO-US30365.
XX
PR   27-SEP-2000; 2000US-235713P.
PR   16-JAN-2001; 2001US-261781P.
PR   19-JUL-2001; 2001US-306605P.
PR   03-AUG-2001; 2001US-310436P.
XX
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI   Feder JN,  Mintier G,  Ramanathan CS,  Hawken DR;
XX
DR   WPI; 2002-435196/46.
DR   P-PSDB; AAU97158.
XX
PT   Novel G protein-coupled receptor, HGPRBMY5 polypeptide, useful for
PT   treatment of AIDS, allergies, asthma, atherosclerosis, ulcerative
PT   colitis, atopic dermatitis, diabetes mellitus, glomerulonephritis,
PT   osteoarthritis -
XX
PS   Claim 1; Fig 1; 148pp; English.
XX
CC   The present invention relates to the isolation of a novel human
CC   G-protein coupled receptor (GPCR) (HGPRBMY5), and the polynucleotide
CC   sequence encoding it. The HGPRBMY5 polypeptide and polynucleotide
CC   are useful for preventing, treating or ameliorating a disease,
CC   disorder or condition related to the colon, brain, ovaries, thymus,
CC   lungs or immune system. They are particularly useful for the
CC   treatment or prevention of cancers, immune disorders, neurological
CC   disorders, and diseases related to the brain, ovaries, thymus or
CC   lungs. The polynucleotide sequence is useful for diagnosing or
CC   determining susceptibility to infections such as bacterial, fungal,
CC   protozoan and viral infections, particularly infections caused by
CC   human immunodeficiency virus (HIV or HIV-2). The present sequence
CC   encodes human GPCR HGPRBMY5.
XX
SQ   Sequence 2214 BP; 638 A; 457 C; 416 G; 703 T; 0 other;

Alignment Scores:
Pred. No.:              3.77e-57      Length:         2214
Score:                  549.00        Matches:        107
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            100.00%       Indels:         0
DB:                     24            Gaps:           0

US-09-930-312-2 (1-107) x ABK51944 (1-2214)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1735 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1794

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1795 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1854

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1855 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1914

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1915 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1974

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1975 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2034

Qy     101 AsnPhePheLysAspLysLeu 107
           |||||||||||||||||||||
Db    2035 AACTTTTTTAAGGACAAGTTG 2055
```

```
RESULT 10
ABL40196
ID   ABL40196 standard; cDNA; 2262 BP.
XX
AC   ABL40196;
XX
DT   23-MAY-2002  (first entry)
XX
DE   Human G protein-coupled receptor TGR17-5 encoding cDNA SEQ ID NO:14.
XX
KW   Human; G protein-coupled receptor; TGR17-5; nootropic; antiinflammatory;
KW   vasotropic; immunomodulator; cytostatic; gene therapy; protein therapy;
KW   neurological; inflammatory; circulatory; degenerative; immune system;
KW   digestive disease; cancer; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..2262
FT                   /*tag= a
FT                   /partial
FT                   /product= "TGR17-5"
FT                   /note= "no stop codon given"
XX
PN   WO200204640-A1.
XX
PD   17-JAN-2002.
XX
PF   06-JUL-2001; 2001WO-JP05878.
XX
PR   07-JUL-2000; 2000JP-0211989.
PR   18-DEC-2000; 2000JP-0383794.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
PI   Moriya T,  Ito T,  Shintani Y,  Miyajima N;
XX
DR   WPI; 2002-179706/23.
DR   P-PSDB; ABB06254.
XX
PT   G-protein coupled receptor protein TGR17 of human origin and DNA
PT   encoding it for diagnosis and treatment of cancer and circulatory and
PT   other diseases associated with its expression -
XX
PS   Claim 7; Page 130-131; 145pp; Japanese.
XX
CC   The present invention describes a human guanine nucleotide binding
CC   protein (G protein) coupled receptor protein designated TGR17, which has
CC   5 variant forms (TGR17-1 to TGR17-5) of the receptor protein having
CC   additional residues at the N-terminal. The TGR17 proteins have nootropic,
CC   antiinflammatory, vasotropic, immunomodulator and cytostatic activities.
CC   The TGR17 polynucleotides and protein can be used in gene therapy and
CC   protein therapy. G-protein coupled receptor proteins are cell membrane
CC   proteins mediating the cellular response to a large variety of signalling
CC   molecules. The TGR17 polynucleotides and proteins can be used in the
CC   diagnosis, treatment and prevention of diseases including neurological,
CC   inflammatory, circulatory, degenerative, immune system and digestive
CC   diseases and cancer. The present sequence encodes human TGR17-5 from the
CC   present invention.
XX
SQ   Sequence 2262 BP; 650 A; 467 C; 421 G; 724 T; 0 other;

Alignment Scores:
Pred. No.:              3.88e-57      Length:          2262
Score:                  549.00        Matches:         107
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            100.00%       Indels:          0
DB:                     24            Gaps:            0

US-09-930-312-2 (1-107) x ABL40196 (1-2262)

Qy      1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1786 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1845

Qy     21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1846 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1905

Qy     41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1906 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1965

Qy     61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1966 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 2025

Qy     81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2026 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2085

Qy    101 AsnPhePheLysAspLysLeu 107
          |||||||||||||||||||||
Db   2086 AACTTTTTTAAGGACAAGTTG 2106

RESULT 11
AAI99584
ID   AAI99584 standard; cDNA; 530 BP.
XX
AC   AAI99584;
XX
DT   04-JAN-2002  (first entry)
XX
DE   Human expressed polynucleotide SEQ ID NO 47.
XX
KW   Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW   immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW   antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW   antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW   antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW   antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; nephrotropic; gene therapy; vaccine;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200155387-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01310.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
```

PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465573/50.
DR P-PSDB; AAM99972.
XX
PT Isolated digestive system associated polypeptide for treating,
PT preventing and/ or prognosing disorders related to the digestive system
PT including digestive system cancers and also for testing and detection
PT e.g. diagnosis -
XX
PS Claim 1; SEQ ID NO 47; 509pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (AAI99548-AAI99604) and proteins
CC (AAM99936-AAM99984) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

```
XX
SQ   Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other;

Alignment Scores:
Pred. No.:              8.34e-37        Length:         530
Score:                  376.00          Matches:        69
Percent Similarity:     83.18%          Conservative:   20
Best Local Similarity:  64.49%          Mismatches:     18
Query Match:            68.49%          Indels:         0
DB:                     22              Gaps:           0

US-09-930-312-2 (1-107) x AAI99584 (1-530)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           |||:::||||||   ||||||:::||||||||||||||||||   :::||||||   |||
Db      32 GGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGT 91

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           :::   :::::::|||:::      ||||||:::|||||||         :::|||:::
Db      92 GTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATC 151

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           :::|||   ||||||||||||||||||||:::||||||:::||||||||||||:::|||
Db     152 CTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTT 211

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           |||||||||   |||||||||   :::||||||||||||   |||:::||||||||:::
Db     212 GTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTA 271

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           |||||||||   ||||||:::||||||||||||||||||||||||||||||||||||||
Db     272 GTGATTTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACA 331

Qy     101 AsnPheLysAspLysLeu 107
                 ||||||:::   :::
Db     332 AGACCATTTAAAGAAATGATT 352

RESULT 12
AAL35644
ID   AAL35644 standard; cDNA; 530 BP.
XX
AC   AAL35644;
XX
DT   08-JAN-2002  (first entry)
XX
DE   Human musculoskeletal system related polynucleotide SEQ ID NO 986.
XX
KW   Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW   antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW   vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW   cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; human; secreted protein;
KW   musculoskeletal system; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155367-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01338.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
```

```
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-451937/48.
DR   P-PSDB; ABB04062.
XX
PT   Isolated polypeptide for treating, preventing and/ or prognosing
PT   disorders related to the musculoskeletal system including
PT   musculoskeletal cancers and also for testing and detection e.g.
PT   diagnosis -
XX
PS   Claim 1; SEQ ID NO 986; 781pp + Sequence Listing; English.
XX
CC   The invention relates to novel genes (AAL34669-AAL37666) and proteins
CC   (ABB03087-ABB04109) associated with the musculoskeletal system useful
CC   for preventing, treating or ameliorating medical conditions e.g. by
CC   protein or gene therapy. The genes are isolated from a range of human
CC   tissues disclosed in the specification. The nucleic acids, proteins,
CC   antibodies and (ant)agonists are useful in the diagnosis, treatment
CC   and prevention of: (a) cancer, e.g. breast and ovarian cancer and
CC   other cancers of the adrenal gland, bone, bone marrow, breast,
CC   gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC   disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC   anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC   multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC   (c) cardiovascular disorders such as myocardial ischaemias; (d) wound
CC   healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC   and (f) infectious diseases such as viral, bacterial, fungal and
CC   parasitic infections.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other;

Alignment Scores:
Pred. No.:              8.34e-37      Length:          530
Score:                  376.00        Matches:         69
Percent Similarity:     83.18%        Conservative:    20
Best Local Similarity:  64.49%        Mismatches:      18
Query Match:            68.49%        Indels:          0
DB:                     22            Gaps:            0

US-09-930-312-2 (1-107) x AAL35644 (1-530)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           |||:::||||||   ||||||:::|||||||||||||||||||   :::|||||||   |||
Db      32 GGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGT 91

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           :::   ::::::|||:::      ||||||:::||||||         :::|||:::
Db      92 GTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATC 151

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           :::|||   ||||||||||||||||||||:::||||||:::||||||||||||:::|||
Db     152 CTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTT 211

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           |||||||||   |||||||||   :::||||||||||||   |||:::|||||||||:::
Db     212 GTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTA 271

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           |||||||||   ||||||:::||||||||||||||||||||||||||||||||||||||
Db     272 GTGATTTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACA 331

Qy     101 AsnPhePheLysAspLysLeu 107
                 ||||||:::   :::
Db     332 AGACCATTTAAAGAAATGATT 352

RESULT 13
ABA06471
ID   ABA06471 standard; cDNA; 530 BP.
XX
AC   ABA06471;
XX
DT   10-JAN-2002  (first entry)
XX
DE   Human cDNA SEQ ID NO: 137.
XX
KW   Human; gene therapy; neural disorder; immune system disorder;
KW   muscular disorder; reproductive disorder; gastrointestinal disorder;
KW   pulmonary disorder; cardiovascular disorder; renal disorder;
KW   proliferative disorder; inflammation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200154474-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01349.
XX
```

PR 31-JAN-2000; 2000US-179065P.
PR 04-FEB-2000; 2000US-180628P.
PR 24-FEB-2000; 2000US-184664P.
PR 02-MAR-2000; 2000US-186350P.
PR 16-MAR-2000; 2000US-189874P.
PR 17-MAR-2000; 2000US-190076P.
PR 18-APR-2000; 2000US-198123P.
PR 19-MAY-2000; 2000US-205515P.
PR 07-JUN-2000; 2000US-209467P.
PR 28-JUN-2000; 2000US-214886P.
PR 30-JUN-2000; 2000US-215135P.
PR 07-JUL-2000; 2000US-216647P.
PR 07-JUL-2000; 2000US-216880P.
PR 11-JUL-2000; 2000US-217487P.
PR 11-JUL-2000; 2000US-217496P.
PR 14-JUL-2000; 2000US-218290P.
PR 26-JUL-2000; 2000US-220963P.
PR 26-JUL-2000; 2000US-220964P.
PR 14-AUG-2000; 2000US-224518P.
PR 14-AUG-2000; 2000US-224519P.
PR 14-AUG-2000; 2000US-225213P.
PR 14-AUG-2000; 2000US-225214P.
PR 14-AUG-2000; 2000US-225266P.
PR 14-AUG-2000; 2000US-225267P.
PR 14-AUG-2000; 2000US-225268P.
PR 14-AUG-2000; 2000US-225270P.
PR 14-AUG-2000; 2000US-225447P.
PR 14-AUG-2000; 2000US-225757P.
PR 14-AUG-2000; 2000US-225758P.
PR 14-AUG-2000; 2000US-225759P.
PR 18-AUG-2000; 2000US-226279P.
PR 22-AUG-2000; 2000US-226681P.
PR 22-AUG-2000; 2000US-226868P.
PR 22-AUG-2000; 2000US-227182P.
PR 23-AUG-2000; 2000US-227009P.
PR 30-AUG-2000; 2000US-228924P.
PR 01-SEP-2000; 2000US-229287P.
PR 01-SEP-2000; 2000US-229343P.
PR 01-SEP-2000; 2000US-229344P.
PR 01-SEP-2000; 2000US-229345P.
PR 05-SEP-2000; 2000US-229509P.
PR 05-SEP-2000; 2000US-229513P.
PR 06-SEP-2000; 2000US-230437P.
PR 06-SEP-2000; 2000US-230438P.
PR 08-SEP-2000; 2000US-231242P.
PR 08-SEP-2000; 2000US-231243P.
PR 08-SEP-2000; 2000US-231244P.
PR 08-SEP-2000; 2000US-231413P.
PR 08-SEP-2000; 2000US-231414P.
PR 08-SEP-2000; 2000US-232080P.
PR 08-SEP-2000; 2000US-232081P.
PR 12-SEP-2000; 2000US-231968P.
PR 14-SEP-2000; 2000US-232397P.
PR 14-SEP-2000; 2000US-232398P.
PR 14-SEP-2000; 2000US-232399P.
PR 14-SEP-2000; 2000US-232400P.
PR 14-SEP-2000; 2000US-232401P.
PR 14-SEP-2000; 2000US-233063P.
PR 14-SEP-2000; 2000US-233064P.
PR 14-SEP-2000; 2000US-233065P.
PR 21-SEP-2000; 2000US-234223P.
PR 21-SEP-2000; 2000US-234274P.
PR 25-SEP-2000; 2000US-234997P.
PR 25-SEP-2000; 2000US-234998P.
PR 26-SEP-2000; 2000US-235484P.
PR 27-SEP-2000; 2000US-235834P.
PR 27-SEP-2000; 2000US-235836P.
PR 29-SEP-2000; 2000US-236327P.
PR 29-SEP-2000; 2000US-236367P.
PR 29-SEP-2000; 2000US-236368P.
PR 29-SEP-2000; 2000US-236369P.
PR 29-SEP-2000; 2000US-236370P.
PR 02-OCT-2000; 2000US-236802P.
PR 02-OCT-2000; 2000US-237037P.
PR 02-OCT-2000; 2000US-237038P.
PR 02-OCT-2000; 2000US-237039P.
PR 02-OCT-2000; 2000US-237040P.
PR 13-OCT-2000; 2000US-239935P.
PR 13-OCT-2000; 2000US-239937P.
PR 20-OCT-2000; 2000US-240960P.
PR 20-OCT-2000; 2000US-241221P.
PR 20-OCT-2000; 2000US-241785P.
PR 20-OCT-2000; 2000US-241786P.
PR 20-OCT-2000; 2000US-241787P.
PR 20-OCT-2000; 2000US-241808P.
PR 20-OCT-2000; 2000US-241809P.
PR 20-OCT-2000; 2000US-241826P.
PR 01-NOV-2000; 2000US-244617P.
PR 08-NOV-2000; 2000US-246474P.
PR 08-NOV-2000; 2000US-246475P.
PR 08-NOV-2000; 2000US-246476P.
PR 08-NOV-2000; 2000US-246477P.
PR 08-NOV-2000; 2000US-246478P.
PR 08-NOV-2000; 2000US-246523P.
PR 08-NOV-2000; 2000US-246524P.
PR 08-NOV-2000; 2000US-246525P.
PR 08-NOV-2000; 2000US-246526P.
PR 08-NOV-2000; 2000US-246527P.
PR 08-NOV-2000; 2000US-246528P.
PR 08-NOV-2000; 2000US-246532P.
PR 08-NOV-2000; 2000US-246609P.
PR 08-NOV-2000; 2000US-246610P.
PR 08-NOV-2000; 2000US-246611P.
PR 08-NOV-2000; 2000US-246613P.
PR 17-NOV-2000; 2000US-249207P.
PR 17-NOV-2000; 2000US-249208P.
PR 17-NOV-2000; 2000US-249209P.
PR 17-NOV-2000; 2000US-249210P.
PR 17-NOV-2000; 2000US-249211P.
PR 17-NOV-2000; 2000US-249212P.
PR 17-NOV-2000; 2000US-249213P.
PR 17-NOV-2000; 2000US-249214P.
PR 17-NOV-2000; 2000US-249215P.
PR 17-NOV-2000; 2000US-249216P.
PR 17-NOV-2000; 2000US-249217P.
PR 17-NOV-2000; 2000US-249218P.
PR 17-NOV-2000; 2000US-249244P.
PR 17-NOV-2000; 2000US-249245P.
PR 17-NOV-2000; 2000US-249264P.
PR 17-NOV-2000; 2000US-249265P.
PR 17-NOV-2000; 2000US-249297P.
PR 17-NOV-2000; 2000US-249299P.
PR 17-NOV-2000; 2000US-249300P.
PR 01-DEC-2000; 2000US-250160P.
PR 01-DEC-2000; 2000US-250391P.
PR 05-DEC-2000; 2000US-251030P.
PR 05-DEC-2000; 2000US-251988P.
PR 05-DEC-2000; 2000US-256719P.
PR 06-DEC-2000; 2000US-251479P.
PR 08-DEC-2000; 2000US-251856P.
PR 08-DEC-2000; 2000US-251868P.
PR 08-DEC-2000; 2000US-251869P.
PR 08-DEC-2000; 2000US-251989P.
PR 08-DEC-2000; 2000US-251990P.
PR 11-DEC-2000; 2000US-254097P.
PR 05-JAN-2001; 2001US-259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-476161/51.
DR P-PSDB; ABB10249.
XX
PT Isolated nucleic acid molecule encoding an inflammation-associated
PT polypeptide is used in preventing, treating or ameliorating a medical

```
PT   condition -
XX
PS   Claim 1; SEQ ID NO: 137; 859pp + Sequence Listing; English.
XX
CC   The present invention provides human cDNAs, proteins and related genomic
CC   DNAs. These can be used in the treatment of neural, immune system,
CC   muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC   renal and proliferative disorders and inflammation. The present sequence
CC   is a cDNA of the invention.
XX
SQ   Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other;

Alignment Scores:
Pred. No.:              8.34e-37        Length:         530
Score:                  376.00          Matches:        69
Percent Similarity:     83.18%          Conservative:   20
Best Local Similarity:  64.49%          Mismatches:     18
Query Match:            68.49%          Indels:         0
DB:                     22              Gaps:           0

US-09-930-312-2 (1-107) x ABA06471 (1-530)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           |||:::||||||   ||||||:::|||||||||||||||||||   :::||||||   |||
Db      32 GGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGT 91

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           :::   ::::::|||:::      ||||||:::||||||         :::|||:::
Db      92 GTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATC 151

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           :::|||   ||||||||||||||||||||:::||||||:::|||||||||||||:::|||
Db     152 CTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTT 211

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           |||||||||   |||||||||   :::|||||||||||||   |||:::||||||||:::
Db     212 GTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTA 271

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           |||||||||   ||||||:::||||||||||||||||||||||||||||||||||||||||
Db     272 GTGATTTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACA 331

Qy     101 AsnPhePheLysAspLysLeu 107
                 |||||||:::   :::
Db     332 AGACCATTTAAAGAAATGATT 352

RESULT 14
AAS28950
ID   AAS28950 standard; cDNA; 530 BP.
XX
AC   AAS28950;
XX
DT   21-NOV-2001  (first entry)
XX
DE   cDNA encoding for human uterine motility-association polypeptide #15.
XX
KW   Human; uterine motility-association disorder; uterus; pregnancy;
KW   labour; menstrual cycle; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155201-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01317.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
```

PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-488777/53.
DR P-PSDB; AAU18108.
XX
PT Isolated polypeptide and nucleic acid molecules for treating,
PT preventing and/or prognosing disorders related to uterine motility
PT e.g. disorders associated with pregnancy and the menstrual cycle -
XX
PS Claim 4; SEQ ID No 25; 524pp; English.
XX
CC The present invention relates to the isolation of novel human
CC uterine motility-association polypeptides (AAU18094-AAU18152),
CC and cDNA and genomic sequences encoding for these polypeptides.
CC The sequences of the invention are useful in the diagnosis,
CC treatment, prevention and/or prognosis of diseases associated
CC with uterine motility such as pregnancy and labour, and menstrual
CC disorders. The polynucleotide sequences of the invention are also
CC useful in gene therapy. AAS28936-AAS28994 represent cDNA sequences
CC encoding for novel human uterine motility-association polypeptides.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other;

Alignment Scores:

Pred. No.:	8.34e-37	Length:	530
Score:	376.00	Matches:	69
Percent Similarity:	83.18%	Conservative:	20
Best Local Similarity:	64.49%	Mismatches:	18
Query Match:	68.49%	Indels:	0
DB:	22	Gaps:	0

US-09-930-312-2 (1-107) x AAS28950 (1-530)

```
Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           |||:::||||||   ||||||:::||||||||||||||||||   :::||||||   |||
Db      32 GGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGT 91

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           :::   ::::::|||:::     ||||||:::||||||          :::|||:::
Db      92 GTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATC 151

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           :::|||   ||||||||||||||||||||:::||||||:::||||||||||||:::|||
Db     152 CTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTT 211

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           |||||||||   |||||||||   :::||||||||||||   |||:::||||||||:::
Db     212 GTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTA 271

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           |||||||||   ||||||:::||||||||||||||||||||||||||||||||||||||
Db     272 GTGATTTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACA 331

Qy     101 AsnPhePheLysAspLysLeu 107
                 ||||||:::   :::
Db     332 AGACCATTTAAAGAAATGATT 352
```

RESULT 15
AAS29573
ID AAS29573 standard; cDNA; 530 BP.
XX
AC AAS29573;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human endocrine polypeptide encoding cDNA SEQ ID No 73.
XX
KW Endocrine protein; human; mouse; rabbit; goat; horse; food additive;
KW cat; dog; chicken; sheep; immunosuppressive; antiarthritic; vasotropic;
KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
KW ophthalmological; vulnerary; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ss;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; food preservative;
KW tissue regeneration; anti-infertility.
XX
OS Homo sapiens.
XX
PN WO200155364-A2.
XX

PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01308.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-451936/48.

```
DR   P-PSDB; AAU18344.
XX
PT   Isolated polypeptide for treating, preventing and/ or prognosing
PT   disorders of the endocrine system such as reproductive disorders,
PT   endocrine cancers and also for testing and detection e.g. diagnosis -
XX
PS   Claim 1; SEQ ID No 73; 604pp; English.
XX
CC   Sequences AAS29511-AAS29736 represent cDNA molecules, which encode the
CC   endocrine polypeptides of the invention. Endocrine polypeptides and their
CC   associated polynucleotides of the invention are useful in the diagnosis,
CC   treatment and prevention of various types of disorders in e.g. humans,
CC   mice, rabbits, goats, horses, cats, dogs, chickens or sheep. A
CC   pathological condition can be determined by determining the presence or
CC   absence of a mutation in an endocrine polynucleotide. The treatable
CC   disorders include autoimmune diseases such as rheumatoid arthritis,
CC   hyperproliferative disorders such as neoplasms of the breast or liver,
CC   cardiovascular disorders such as cardiac arrest, cerebrovascular
CC   disorders such as cerebral ischaemia, nervous system disorders such as
CC   Alzheimer's disease, infections caused by bacteria, viruses and fungi,
CC   ocular disorders such as corneal infection, endocrine disorders such as
CC   premature labour and infertility, gastrointestinal disorders such as
CC   Crohn's disease, renal disorders such as glomerulonephritis and
CC   respiratory disorders such as asthma. The polypeptides can also be used
CC   to aid wound healing, to prevent skin aging due to sunburn, to maintain
CC   organs before transplantation, to regenerate tissues and in chemotaxis.
CC   The polypeptides can also be used as a food additive or preservative to
CC   increase or decrease storage capabilities.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO

Alignment Scores:
Pred. No.:              8.34e-37      Length:          530
Score:                  376.00        Matches:         69
Percent Similarity:     83.18%        Conservative:    20
Best Local Similarity:  64.49%        Mismatches:      18
Query Match:            68.49%        Indels:          0
DB:                     22            Gaps:            0

US-09-930-312-2 (1-107) x AAS29573 (1-530)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           |||:::|||||||   |||||||:::||||||||||||||||||   :::|||||||   |||
Db      32 GGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGT 91

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           :::   :::::::|||:::      ||||||:::|||||||         :::|||:::
Db      92 GTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATC 151

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           :::|||   ||||||||||||||||||||:::|||||||:::||||||||||||:::|||
Db     152 CTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTT 211

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           |||||||||   |||||||||   :::||||||||||||   |||:::||||||||:::
Db     212 GTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTA 271

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           |||||||||   ||||||:::||||||||||||||||||||||||||||||||||||||
Db     272 GTGATTTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACA 331

Qy     101 AsnPhePheLysAspLysLeu 107
                 |||||||:::   :::
Db     332 AGACCATTTAAAGAAATGATT 352


Search completed: May 26, 2003, 20:11:36
Job time : 241 secs
```

GenCore version 5.1.4_p5_4578

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 26, 2003, 20:05:47 ; Search time 52 Seconds
(without alignments)
631.046 Million cell updates/sec

Title: US-09-930-312-2
Perfect score: 549
Sequence: 1 GVNLLAFLIIVFSYITMFCS.....NSALNPILYTLTTNFFKDKL 107

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09930312/runat_22052003_114614_27092/app_query.fasta_1.263
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09930312_@CGN_1_1_40_@runat_22052003_114614_27092 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	138.5	25.2	1831	1	US-08-101-435-3	Sequence 3, Appli
2	138.5	25.2	2022	4	US-07-757-342D-9	Sequence 9, Appli
3	138.5	25.2	2987	4	US-07-757-342D-1	Sequence 1, Appli
c 4	125	22.8	1298	4	US-08-795-876-32	Sequence 32, Appl
5	125	22.8	1300	4	US-08-795-876-31	Sequence 31, Appl
6	125	22.8	1300	4	US-08-795-876-36	Sequence 36, Appl
c 7	125	22.8	1300	4	US-08-795-876-37	Sequence 37, Appl
8	125	22.8	2179	1	US-08-487-886-1	Sequence 1, Appli
9	125	22.8	2179	2	US-08-531-070A-1	Sequence 1, Appli
10	125	22.8	2179	3	US-08-482-855-1	Sequence 1, Appli
11	125	22.8	2179	4	US-08-474-986-1	Sequence 1, Appli
12	123	22.4	981	1	US-08-349-696-20	Sequence 20, Appl
13	123	22.4	981	1	US-08-233-009-20	Sequence 20, Appl
14	123	22.4	981	1	US-08-560-231-20	Sequence 20, Appl
15	123	22.4	981	4	US-09-080-704A-20	Sequence 20, Appl
16	123	22.4	2180	1	US-07-918-314-3	Sequence 3, Appli
17	123	22.4	2900	1	US-07-918-314-5	Sequence 5, Appli
18	120.5	21.9	4417	4	US-07-741-453A-57	Sequence 57, Appl
19	119.5	21.8	957	1	US-08-349-696-26	Sequence 26, Appl
20	119.5	21.8	957	1	US-08-233-009-26	Sequence 26, Appl
21	119.5	21.8	957	1	US-08-560-231-26	Sequence 26, Appl
22	119.5	21.8	957	4	US-09-080-704A-26	Sequence 26, Appl
23	119.5	21.8	3710	4	US-07-741-453A-62	Sequence 62, Appl
24	116	21.1	1859	1	US-08-293-563-1	Sequence 1, Appli
25	115	20.9	999	1	US-08-349-696-24	Sequence 24, Appl
26	115	20.9	999	1	US-08-233-009-24	Sequence 24, Appl
27	115	20.9	999	1	US-08-560-231-24	Sequence 24, Appl
28	115	20.9	999	4	US-09-080-704A-24	Sequence 24, Appl
29	109	19.9	473	2	US-08-866-757-3	Sequence 3, Appli
30	109	19.9	473	4	US-09-153-593-3	Sequence 3, Appli
31	109	19.9	4203	2	US-08-866-757-1	Sequence 1, Appli
32	109	19.9	4203	4	US-09-153-593-1	Sequence 1, Appli
33	108	19.7	1205	1	US-08-417-103-13	Sequence 13, Appl
34	108	19.7	1265	1	US-07-816-283-3	Sequence 3, Appli
35	108	19.7	1265	1	US-08-417-103-3	Sequence 3, Appli
36	108	19.7	1634	1	US-07-816-283-1	Sequence 1, Appli
37	108	19.7	1634	1	US-08-417-103-1	Sequence 1, Appli
38	105	19.1	1149	4	US-09-262-477-1	Sequence 1, Appli
39	103.5	18.9	2862	1	US-08-148-209A-1	Sequence 1, Appli
40	103.5	18.9	2884	1	US-08-148-209A-5	Sequence 5, Appli
41	102.5	18.7	1639	1	US-08-334-698-5	Sequence 5, Appli
42	102.5	18.7	1639	1	US-08-228-932-5	Sequence 5, Appli
43	102.5	18.7	1639	1	US-08-468-939-5	Sequence 5, Appli
44	102.5	18.7	1639	2	US-08-406-855A-5	Sequence 5, Appli
45	102.5	18.7	1639	2	US-08-722-190-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-101-435-3
; Sequence 3, Application US/08101435
; Patent No. 5441883
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Zhou, Qun-yong
; TITLE OF INVENTION: A No. 5441883el Adenosine Receptor and Uses
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/101,435
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/847,563
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 91,708
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317

```
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1831 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  316..1276
US-08-101-435-3

Alignment Scores:
Pred. No.:              3.87e-08      Length:         1831
Score:                  138.50        Matches:        29
Percent Similarity:     52.58%        Conservative:   22
Best Local Similarity:  29.90%        Mismatches:     41
Query Match:            25.23%        Indels:         5
DB:                     1             Gaps:           2

US-09-930-312-2 (1-107) x US-08-101-435-3 (1-1831)

Qy      4 LeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSerIleGlnLys 23
          |||:::      :::::::::         |||:::   :::|||      |||:::
Db    880 CTCATCCCCCTGGTTGTCATGTGCATCATCTATCTGGACATCTTCTACATCATCCGAAAC 939

Qy     24 ThrAlaLeuGlnThr----------ThrGluValArgAsnCysPheGlyArgGluVal 39
                  |||                  |||   |||      :::||||||||
Db    940 AAACTCAGTCAGAATCTGACTGGCTTCAGAGAGACGCGTGCATTTTACGGTCGGGAGTTC 999

Qy     40 AlaValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProVal 59
             |||         |||   ::::::|||      |||:::||||||:::|||:::
Db   1000 AAGACCGCTAAGTCCCTGTTTCTGGTTCTCTTCTTGTTTGCCTTGTGCTGGCTGCCTTTG 1059

Qy     60 PheValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrp 79
          ::::::      :::|||   |||   |||:::||||||:::   :::
Db   1060 TCCATCATCAATTTTGTTTCCTACTTTAATGTGAAGATACCAGAG---ATTGCAATGTGC 1116

Qy     80 IleValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyr 96
          :::   |||              ||||||   :::||||||||||:::|||
Db   1117 CTGGGCATCCTGTTGTCCCATGCGAACTCCATGATGAACCCTATTGTCTAC 1167

RESULT 2
US-07-757-342D-9
; Sequence 9, Application US/07757342D
; Patent No. 6218509
;   GENERAL INFORMATION:
;        APPLICANT: IGARASHI, Masao
;                   MINEGISHI, Takashi
;                   NAKAMURA, Kazuto
;        TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
;        NUMBER OF SEQUENCES: 10
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
;                        CUSHMAN
;             STREET: 130 Water Street
;             CITY: Boston
;             STATE: Massachusetts
;             COUNTRY: US
;             ZIP: 02109
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/07/757,342D
;             FILING DATE: 10-Sep-1991
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BUCKLEY, Linda M.
;             REGISTRATION NUMBER: 31003
;             REFERENCE/DOCKET NUMBER: 41226
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617)523-3400
;             TELEFAX: (617)523-6440
;             TELEX: 200291 STRE UR
;   INFORMATION FOR SEQ ID NO: 9:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2022 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA to mRNA
;        FEATURE:
;             NAME/KEY:  CDS
;             LOCATION:  1..2022
;        SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-07-757-342D-9

Alignment Scores:
Pred. No.:              4.41e-08      Length:         2022
Score:                  138.50        Matches:        29
Percent Similarity:     51.92%        Conservative:   25
Best Local Similarity:  27.88%        Mismatches:     43
Query Match:            25.23%        Indels:         7
DB:                     4             Gaps:           2

US-09-930-312-2 (1-107) x US-07-757-342D-9 (1-2022)

Qy      2 ValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSerIle 21
          :::|||:::::::|||||   ||||||         ||||||   ::::::   ::::::
Db   1525 CTCAATGTGGTGGCCTTCTTCATAATTTGTGCTTGCTACATTAAAATTTATTTTGCAGTT 1584

Qy     22 GlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAlaVal 41
          :::         |||      |||                      ::::::      :::
Db   1585 CGAAACCCAGAATTAATGGCTACC------------------AATAAAGATACAAAGATT 1626

Qy     42 AlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPheVal 61
          |||   :::          :::::|||:::|||      |||      |||:::
Db   1627 GCTAAGAAAATGGCAATCCTCATCTTCACCGATTTCACCTGCATGGCACCTATCTCTTTT 1686

Qy     62 ValLysIleLeuSerLeuPheArgValGlu---IleProAspThrMetThrSerTrpIle 80
               |||   :::   |||:::|||      |||      |||   :::      :::
Db   1687 TTTGCCATCTCAGCTGCCTTCAAAGTACCTCTTATCACAGTAACCAACTCTAAAGTTTTA 1746

Qy     81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
          ::::::   |||   |||:::||||||      ||||||   ||||||   ::::  |||
Db   1747 CTGGTTCTTTTTTATCCCATCAATTCTTGTGCCAATCCATTTCTGTATGCAATATTCACT 1806

Qy    101 AsnPhePheLys 104
               |||:::
Db   1807 AAGACATTCCAA 1818

RESULT 3
US-07-757-342D-1
; Sequence 1, Application US/07757342D
; Patent No. 6218509
;   GENERAL INFORMATION:
;        APPLICANT: IGARASHI, Masao
;                   MINEGISHI, Takashi
;                   NAKAMURA, Kazuto
;        TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
;        NUMBER OF SEQUENCES: 10
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
;                        CUSHMAN
;             STREET: 130 Water Street
;             CITY: Boston
;             STATE: Massachusetts
;             COUNTRY: US
;             ZIP: 02109
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;          CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/07/757,342D
;            FILING DATE: 10-Sep-1991
;            CLASSIFICATION: <Unknown>
;          ATTORNEY/AGENT INFORMATION:
;            NAME: BUCKLEY, Linda M.
;            REGISTRATION NUMBER: 31003
;            REFERENCE/DOCKET NUMBER: 41226
;          TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)523-3400
;            TELEFAX: (617)523-6440
;            TELEX: 200291 STRE UR
;     INFORMATION FOR SEQ ID NO: 1:
;          SEQUENCE CHARACTERISTICS:
;            LENGTH: 2987 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;          MOLECULE TYPE: cDNA to mRNA
;          FEATURE:
;            NAME/KEY:  CDS
;            LOCATION:  1..2097
;          SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-07-757-342D-1

Alignment Scores:
Pred. No.:              7.35e-08        Length:         2987
Score:                  138.50          Matches:        29
Percent Similarity:     51.92%          Conservative:   25
Best Local Similarity:  27.88%          Mismatches:     43
Query Match:            25.23%          Indels:         7
DB:                     4               Gaps:           2

US-09-930-312-2 (1-107) x US-07-757-342D-1 (1-2987)

Qy       2 ValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSerIle 21
           :::|||:::::|||||   ||||||          ||||||   ::::::   ::::::
Db    1600 CTCAATGTGGTGGCCTTCTTCATAATTTGTGCTTGCTACATTAAAATTTATTTTGCAGTT 1659

Qy      22 GlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAlaVal 41
           :::         |||      |||                      ::::::      :::
Db    1660 CGAAACCCAGAATTAATGGCTACC-----------------AATAAAGATACAAAGATT 1701

Qy      42 AlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPheVal 61
           |||   :::         ::::::|||:::|||      |||      |||:::
Db    1702 GCTAAGAAAATGGCAATCCTCATCTTCACCGATTTCACCTGCATGGCACCTATCTCTTTT 1761

Qy      62 ValLysIleLeuSerLeuPheArgValGlu---IleProAspThrMetThrSerTrpIle 80
             |||   :::   |||:::|||      |||      |||   :::      :::
Db    1762 TTTGCCATCTCAGCTGCCTTCAAAGTACCTCTTATCACAGTAACCAACTCTAAAGTTTTA 1821

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           ::::::   |||   |||:::||||||      ||||||   ||||||   ::   |||
Db    1822 CTGGTTCTTTTTTATCCCATCAATTCTTGTGCCAATCCATTTCTGTATGCAATATTCACT 1881

Qy     101 AsnPheLys 104
             |||:::
Db    1882 AAGACATTCCAA 1893

RESULT 4
US-08-795-876-32/c
; Sequence 32, Application US/08795876
; Patent No. 6403305
;  GENERAL INFORMATION:
;    APPLICANT:  Gershengorn, Marvin C.
;    APPLICANT:  Geras-Raaka, Elizabeth
;    APPLICANT:  Nussenzveig, Daniel R.
;    TITLE OF INVENTION:  STRATEGY TO CLONE DRUGS FOR G PROTEIN
;    TITLE OF INVENTION:  COUPLED RECEPTORS
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON, HARGRAVE, DEVANS & DOYLE LLP
;      STREET:  Clinton Square, P.O. Box 1051
;      CITY:  Rochester
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  14603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/795,876
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BRAMAN, SUSAN J.
;      REGISTRATION NUMBER:  34,103
;      REFERENCE/DOCKET NUMBER:  19603/1280
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  716-263-1636
;      TELEFAX:  716-263-1600
;  INFORMATION FOR SEQ ID NO:  32:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1298 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-795-876-32

Alignment Scores:
Pred. No.:              1.21e-06        Length:         1298
Score:                  125.00          Matches:        28
Percent Similarity:     52.83%          Conservative:   28
Best Local Similarity:  26.42%          Mismatches:     42
Query Match:            22.77%          Indels:         8
DB:                     4               Gaps:           3

US-09-930-312-2 (1-107) x US-08-795-876-32 (1-1298)

Qy       2 ValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSerIle 21
           :::|||:::|||||||||:::::::|||          ||||||   ::::::   ::::::
Db     486 CTCAATGTCCTGGCCTTTGTGGTCATCTGTGGCTGCTATATCCACATCTACCTCACAGTG 427

Qy      22 GlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAlaVal 41
           :::         :::   ::::::                      :::      :::
Db     426 CGGAACCCCAACATCGTGTCCTCC-----------------TCTAGTGACACCAGGATC 385

Qy      42 AlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPheVal 61
           |||   |||          :::::::|||:::|||   :::|||      |||:::
Db     384 GCCAAGCGCATGGCCATGCTCATCTTCACTGACTTCCTCTGCATGGCACCCATTTCTTTC 325

Qy      62 ValLysIleLeuSerLeuPheArgValGlu---IleProAspThrMetThrSerTrpIle 80
             |||   :::.      :::|||      |||      :::             :::
Db     324 TTTGCCATTTCTGCCTCCCTCAAGGTGCCCCTCATCACTGTGTCCAAAGCAAAGATTCTG 265

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeu---Thr 99
           ::::::   |||   |||:::||||||      ||||||   ||||||   :::   |||
Db     264 CTGGTTCTGTTTCACCCCATCAACTCCTGTGCCAACCCCTTCCTCTATGCCATCTTTACC 205

Qy     100 ThrAsnPhePheLysAsp 105
             ||||||   :::|||
Db     204 AAAAACTTTCGCAGAGAT 187

RESULT 5
US-08-795-876-31
; Sequence 31, Application US/08795876
; Patent No. 6403305
;  GENERAL INFORMATION:
;    APPLICANT:  Gershengorn, Marvin C.
;    APPLICANT:  Geras-Raaka, Elizabeth
;    APPLICANT:  Nussenzveig, Daniel R.
;    TITLE OF INVENTION:  STRATEGY TO CLONE DRUGS FOR G PROTEIN
```

```
;    TITLE OF INVENTION:  COUPLED RECEPTORS
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON, HARGRAVE, DEVANS & DOYLE LLP
;      STREET:  Clinton Square, P.O. Box 1051
;      CITY:  Rochester
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  14603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/795,876
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BRAMAN, SUSAN J.
;      REGISTRATION NUMBER:  34,103
;      REFERENCE/DOCKET NUMBER:  19603/1280
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  716-263-1636
;      TELEFAX:  716-263-1600
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1300 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-795-876-31

Alignment Scores:
Pred. No.:              1.22e-06      Length:         1300
Score:                  125.00        Matches:        28
Percent Similarity:     52.83%        Conservative:   28
Best Local Similarity:  26.42%        Mismatches:     42
Query Match:            22.77%        Indels:         8
DB:                     4             Gaps:           3

US-09-930-312-2 (1-107) x US-08-795-876-31 (1-1300)

Qy       2 ValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSerIle 21
           :::|||:::|||||||||||:::::|||         ||||||   ::::::   ::::::
Db     819 CTCAATGTCCTGGCCTTTGTGGTCATCTGTGGCTGCTATATCCACATCTACCTCACAGTG 878

Qy      22 GlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAlaVal 41
           :::          :::   ::::::                  :::      :::
Db     879 CGGAACCCCAACATCGTGTCCTCC----------------TCTAGTGACACCAGGATC 920

Qy      42 AlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPheVal 61
           |||   |||         ::::::|||:::|||   :::|||      |||:::
Db     921 GCCAAGCGCATGGCCATGCTCATCTTCACTGACTTCCTCTGCATGGCACCCATTTCTTTC 980

Qy      62 ValLysIleLeuSerLeuPheArgValGlu---IleProAspThrMetThrSerTrpIle 80
             |||   :::      :::|||      |||      :::             :::
Db     981 TTTGCCATTTCTGCCTCCCTCAAGGTGCCCCTCATCACTGTGTCCAAAGCAAAGATTCTG 1040

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeu---Thr 99
           ::::::   |||   |||:::||||||      ||||||   ||||||   :::   |||
Db    1041 CTGGTTCTGTTTCACCCCATCAACTCCTGTGCCAACCCCTTCCTCTATGCCATCTTTACC 1100

Qy     100 ThrAsnPhePheLysAsp 105
             ||||||   :::|||
Db    1101 AAAAACTTTCGCAGAGAT 1118

RESULT 6
US-08-795-876-36
; Sequence 36, Application US/08795876
; Patent No. 6403305
;  GENERAL INFORMATION:
;    APPLICANT:  Gershengorn, Marvin C.
;    APPLICANT:  Geras-Raaka, Elizabeth
;    APPLICANT:  Nussenzveig, Daniel R.
;    TITLE OF INVENTION:  STRATEGY TO CLONE DRUGS FOR G PROTEIN
;    TITLE OF INVENTION:  COUPLED RECEPTORS
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON, HARGRAVE, DEVANS & DOYLE LLP
;      STREET:  Clinton Square, P.O. Box 1051
;      CITY:  Rochester
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  14603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/795,876
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BRAMAN, SUSAN J.
;      REGISTRATION NUMBER:  34,103
;      REFERENCE/DOCKET NUMBER:  19603/1280
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  716-263-1636
;      TELEFAX:  716-263-1600
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1300 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-795-876-36

Alignment Scores:
Pred. No.:              1.22e-06      Length:         1300
Score:                  125.00        Matches:        28
Percent Similarity:     52.83%        Conservative:   28
Best Local Similarity:  26.42%        Mismatches:     42
Query Match:            22.77%        Indels:         8
DB:                     4             Gaps:           3

US-09-930-312-2 (1-107) x US-08-795-876-36 (1-1300)

Qy       2 ValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSerIle 21
           :::|||:::|||||||||:::::|||         ||||||   ::::::   ::::::
Db     819 CTCAATGTCCTGGCCTTTGTGGTCATCTGTGGCTGCTATATCCACATCTACCTCACAGTG 878

Qy      22 GlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAlaVal 41
           :::          :::   ::::::                  :::      :::
Db     879 CGGAACCCCAACATCGTGTCCTCC----------------TCTAGTGACACCAGGATC 920

Qy      42 AlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPheVal 61
           |||   |||         ::::::|||:::|||   :::|||      |||:::
Db     921 GCCAAGCGCATGGCCATGCTCATCTTCACTGACTTCCTCTGCATGGCACCCATTTCTTTC 980

Qy      62 ValLysIleLeuSerLeuPheArgValGlu---IleProAspThrMetThrSerTrpIle 80
             |||   :::      :::|||      |||      :::             :::
Db     981 TTTGCCATTTCTGCCTCCCTCAAGGTGCCCCTCATCACTGTGTCCAAAGCAAAGATTCTG 1040

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeu---Thr 99
           ::::::   |||   |||:::||||||      ||||||   ||||||   :::   |||
Db    1041 CTGGTTCTGTTTCACCCCATCAACTCCTGTGCCAACCCCTTCCTCTATGCCATCTTTACC 1100

Qy     100 ThrAsnPhePheLysAsp 105
             ||||||   :::|||
Db    1101 AAAAACTTTCGCAGAGAT 1118

RESULT 7
```

```
US-08-795-876-37/c
; Sequence 37, Application US/08795876
; Patent No. 6403305
;  GENERAL INFORMATION:
;    APPLICANT:  Gershengorn, Marvin C.
;    APPLICANT:  Geras-Raaka, Elizabeth
;    APPLICANT:  Nussenzveig, Daniel R.
;    TITLE OF INVENTION:  STRATEGY TO CLONE DRUGS FOR G PROTEIN
;    TITLE OF INVENTION:  COUPLED RECEPTORS
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON, HARGRAVE, DEVANS & DOYLE LLP
;      STREET:  Clinton Square, P.O. Box 1051
;      CITY:  Rochester
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  14603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/795,876
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BRAMAN, SUSAN J.
;      REGISTRATION NUMBER:  34,103
;      REFERENCE/DOCKET NUMBER:  19603/1280
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  716-263-1636
;      TELEFAX:  716-263-1600
;  INFORMATION FOR SEQ ID NO:  37:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1300 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-795-876-37

Alignment Scores:
Pred. No.:              1.22e-06        Length:         1300
Score:                  125.00          Matches:        28
Percent Similarity:     52.83%          Conservative:   28
Best Local Similarity:  26.42%          Mismatches:     42
Query Match:            22.77%          Indels:         8
DB:                     4               Gaps:           3

US-09-930-312-2 (1-107) x US-08-795-876-37 (1-1300)

Qy      2 ValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSerIle 21
          :::|||:::|||||||||||:::::|||         ||||||   ::::::   ::::::
Db    486 CTCAATGTCCTGGCCTTTGTGGTCATCTGTGGCTGCTATATCCACATCTACCTCACAGTG 427

Qy     22 GlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAlaVal 41
          :::         :::   ::::::                  :::      :::
Db    426 CGGAACCCCAACATCGTGTCCTCC----------------TCTAGTGACACCAGGATC 385

Qy     42 AlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPheVal 61
          |||   |||         :::::::|||:::|||   :::|||      |||:::
Db    384 GCCAAGCGCATGGCCATGCTCATCTTCACTGACTTCCTCTGCATGGCACCCATTTCTTTC 325

Qy     62 ValLysIleLeuSerLeuPheArgValGlu---IleProAspThrMetThrSerTrpIle 80
                |||   :::      :::|||      |||      :::            :::
Db    324 TTTGCCATTTCTGCCTCCCTCAAGGTGCCCCTCATCACTGTGTCCAAAGCAAAGATTCTG 265

Qy     81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeu---Thr 99
          ::::::   |||   |||:::||||||     ||||||   ||||||   :::   |||
Db    264 CTGGTTCTGTTTCACCCCATCAACTCCTGTGCCAACCCCTTCCTCTATGCCATCTTTACC 205

Qy    100 ThrAsnPhePheLysAsp 105
```

```
                |||||   :::|||
Db    204 AAAAACTTTCGCAGAGAT 187

RESULT 8
US-08-487-886-1
; Sequence 1, Application US/08487886
; Patent No. 5744448
;  GENERAL INFORMATION:
;    APPLICANT:  Kelton, Christie Ann
;    APPLICANT:  Schweickhardt, Rene Lynn
;    APPLICANT:  Cheng, Shirley Vui Yen
;    APPLICANT:  Nugent, No. 5744448een Patrice
;    TITLE OF INVENTION:  Human Follicle Stimulating
;    TITLE OF INVENTION:  Hormone Receptor
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Stephan P. Williams,
;      ADDRESSEE:  Ares-Serono, Inc.
;      STREET:  Exchange Place, 37th floor
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" diskette, 1.44 MB, high density
;      COMPUTER:  IBM PS/2, model 55 SX
;      OPERATING SYSTEM:  MS-DOS version 4.0
;      SOFTWARE:  VAX/VMS Mass11 via Kermit to IBM MS-DOS
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,886
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/670,085
;      FILING DATE:  15-MAR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams, Stephan P.
;      REGISTRATION NUMBER:  28546
;      REFERENCE/DOCKET NUMBER:  US/252
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 723-1300
;      TELEFAX:  (617) 723-8923
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2179
;      TYPE:  Nucleic acid
;      STRANDEDNESS:  Double
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  cDNA to mRNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;      TISSUE TYPE:  Testis
;    IMMEDIATE SOURCE:
;      LIBRARY:  lgt11 cDNA library, ClonTech #HL1010b
;      CLONE:  pHFSHR11-11, pHFSHR15-6
;    FEATURE:
;      NAME/KEY:  protein coding region
;      LOCATION:  75 to 2159
US-08-487-886-1

Alignment Scores:
Pred. No.:              2.39e-06        Length:         2179
Score:                  125.00          Matches:        28
Percent Similarity:     52.83%          Conservative:   28
Best Local Similarity:  26.42%          Mismatches:     42
Query Match:            22.77%          Indels:         8
DB:                     1               Gaps:           3

US-09-930-312-2 (1-107) x US-08-487-886-1 (1-2179)

Qy      2 ValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSerIle 21
          :::|||:::|||||||||||:::::|||         ||||||   ::::::   ::::::
Db   1683 CTCAATGTCCTGGCCTTTGTGGTCATCTGTGGCTGCTATATCCACATCTACCTCACAGTG 1742
```

```
Qy      22 GlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAlaVal 41
           :::         :::   ::::::                          :::      :::
Db    1743 CGGAACCCCAACATCGTGTCCTCC------------------TCTAGTGACACCAGGATC 1784

Qy      42 AlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPheVal 61
           |||   |||          :::::::|||:::|||   :::|||     |||:::
Db    1785 GCCAAGCGCATGGCCATGCTCATCTTCACTGACTTCCTCTGCATGGCACCCATTTCTTTC 1844

Qy      62 ValLysIleLeuSerLeuPheArgValGlu---IleProAspThrMetThrSerTrpIle 80
             |||   :::      :::|||      |||      :::             :::
Db    1845 TTTGCCATTTCTGCCTCCCTCAAGGTGCCCCTCATCACTGTGTCCAAAGCAAAGATTCTG 1904

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeu---Thr 99
           ::::::   |||   |||:::||||||    ||||||   ||||||   :::   |||
Db    1905 CTGGTTCTGTTTCACCCCATCAACTCCTGTGCCAACCCCTTCCTCTATGCCATCTTTACC 1964

Qy     100 ThrAsnPhePheLysAsp 105
           ||||||   :::|||
Db    1965 AAAAACTTTCGCAGAGAT 1982

RESULT 9
US-08-531-070A-1
; Sequence 1, Application US/08531070A
; Patent No. 5851768
;  GENERAL INFORMATION:
;    APPLICANT:  de la Chapelle, Albert
;    APPLICANT:  Aittomaki, Kristiina
;    APPLICANT:  Huhtaniemi, Ilpo
;    TITLE OF INVENTION:  Method For Diagnosis Of Ovarian Dysgenesis
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/531,070A
;      FILING DATE:  20-SEP-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gass, David A.
;      REGISTRATION NUMBER:  38,153
;      REFERENCE/DOCKET NUMBER:  28113/32879
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2179 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-531-070A-1

Alignment Scores:
Pred. No.:              2.39e-06        Length:         2179
Score:                  125.00          Matches:        28
Percent Similarity:     52.83%          Conservative:   28
Best Local Similarity:  26.42%          Mismatches:     42
Query Match:            22.77%          Indels:         8
DB:                     2               Gaps:           3

US-09-930-312-2 (1-107) x US-08-531-070A-1 (1-2179)
```

```
Qy       2 ValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSerIle 21
           :::|||:::|||||||||:::::::|||         ||||||   ::::::   ::::::
Db    1683 CTCAATGTCCTGGCCTTTGTGGTCATCTGTGGCTGCTATATCCACATCTACCTCACAGTG 1742

Qy      22 GlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAlaVal 41
           :::         :::   ::::::                          :::      :::
Db    1743 CGGAACCCCAACATCGTGTCCTCC------------------TCTAGTGACACCAGGATC 1784

Qy      42 AlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPheVal 61
           |||   |||          :::::::|||:::|||   :::|||     |||:::
Db    1785 GCCAAGCGCATGGCCATGCTCATCTTCACTGACTTCCTCTGCATGGCACCCATTTCTTTC 1844

Qy      62 ValLysIleLeuSerLeuPheArgValGlu---IleProAspThrMetThrSerTrpIle 80
             |||   :::      :::|||      |||      :::             :::
Db    1845 TTTGCCATTTCTGCCTCCCTCAAGGTGCCCCTCATCACTGTGTCCAAAGCAAAGATTCTG 1904

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeu---Thr 99
           ::::::   |||   |||:::||||||    ||||||   ||||||   :::   |||
Db    1905 CTGGTTCTGTTTCACCCCATCAACTCCTGTGCCAACCCCTTCCTCTATGCCATCTTTACC 1964

Qy     100 ThrAsnPhePheLysAsp 105
           ||||||   :::|||
Db    1965 AAAAACTTTCGCAGAGAT 1982

RESULT 10
US-08-482-855-1
; Sequence 1, Application US/08482855
; Patent No. 6121016
;  GENERAL INFORMATION:
;    APPLICANT:  Kelton, Christie Ann
;    APPLICANT:  Schweickhardt, Rene Lynn
;    APPLICANT:  Cheng, Shirley Vui Yen
;    APPLICANT:  Nugent, No. 6121016een Patrice
;    TITLE OF INVENTION:  Human Follicle Stimulating
;    TITLE OF INVENTION:  Hormone Receptor
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Stephan P. Williams,
;      ADDRESSEE:  Ares-Serono, Inc.
;      STREET:  Exchange Place, 37th floor
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" diskette, 1.44 MB, high density
;      COMPUTER:  IBM PS/2, model 55 SX
;      OPERATING SYSTEM:  MS-DOS version 4.0
;      SOFTWARE:  VAX/VMS Mass11 via Kermit to IBM MS-DOS
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/482,855
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/670,085
;      FILING DATE:  15-MAR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams, Stephan P.
;      REGISTRATION NUMBER:  28546
;      REFERENCE/DOCKET NUMBER:  US/252
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 723-1300
;      TELEFAX:  (617) 723-8923
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2179
;      TYPE:  Nucleic acid
;      STRANDEDNESS:  Double
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  cDNA to mRNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
```

```
;      TISSUE TYPE:  Testis
;    IMMEDIATE SOURCE:
;      LIBRARY:  lgt11 cDNA library, ClonTech #HL1010b
;      CLONE:  pHFSHR11-11, pHFSHR15-6
;    FEATURE:
;      NAME/KEY:  protein coding region
;      LOCATION:  75 to 2159
US-08-482-855-1

Alignment Scores:
Pred. No.:              2.39e-06        Length:         2179
Score:                  125.00          Matches:        28
Percent Similarity:     52.83%          Conservative:   28
Best Local Similarity:  26.42%          Mismatches:     42
Query Match:            22.77%          Indels:         8
DB:                     3               Gaps:           3

US-09-930-312-2 (1-107) x US-08-482-855-1 (1-2179)

Qy      2 ValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSerIle 21
          :::|||:::|||||||||:::::::|||         ||||||   ::::::   ::::::
Db   1683 CTCAATGTCCTGGCCTTTGTGGTCATCTGTGGCTGCTATATCCACATCTACCTCACAGTG 1742

Qy     22 GlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAlaVal 41
          :::        :::   ::::::                  :::       :::
Db   1743 CGGAACCCCAACATCGTGTCCTCC-------------TCTAGTGACACCAGGATC 1784

Qy     42 AlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPheVal 61
          |||   |||         ::::::|||:::|||   ::::|||      |||:::
Db   1785 GCCAAGCGCATGGCCATGCTCATCTTCACTGACTTCCTCTGCATGGCACCCATTTCTTTC 1844

Qy     62 ValLysIleLeuSerLeuPheArgValGlu---IleProAspThrMetThrSerTrpIle 80
            |||   :::       :::|||      |||      :::             :::
Db   1845 TTTGCCATTTCTGCCTCCCTCAAGGTGCCCCTCATCACTGTGTCCAAAGCAAAGATTCTG 1904

Qy     81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeu---Thr 99
          ::::::   |||   |||:::||||||      ||||||  ||||||   :::   |||
Db   1905 CTGGTTCTGTTTCACCCCATCAACTCCTGTGCCAACCCCTTCCTCTATGCCATCTTTACC 1964

Qy    100 ThrAsnPhePheLysAsp 105
            ||||||   :::|||
Db   1965 AAAAACTTTCGCAGAGAT 1982

RESULT 11
US-08-474-986-1
; Sequence 1, Application US/08474986
; Patent No. 6372711
;   GENERAL INFORMATION:
;        APPLICANT: Kelton, Christie Ann
;                   Schweickhardt, Rene Lynn
;                   Cheng, Shirley Vui Yen
;                   Nugent, No. 637271leen Patrice
;        TITLE OF INVENTION: Human Follicle Stimulating
;                            Hormone Receptor
;        NUMBER OF SEQUENCES: 2
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Stephan P. Williams,
;                        Ares-Serono, Inc.
;             STREET: Exchange Place, 37th floor
;             CITY: Boston
;             STATE: MA
;             COUNTRY: USA
;             ZIP: 02109
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
;             COMPUTER: IBM PS/2, model 55 SX
;             OPERATING SYSTEM: MS-DOS version 4.0
;             SOFTWARE: VAX/VMS Mass11 via Kermit to IBM MS-DOS
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/474,986
;             FILING DATE: 07-Jun-1995
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 07/670,085
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Williams, Stephan P.
;             REGISTRATION NUMBER: 28546
;             REFERENCE/DOCKET NUMBER: US/252
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617) 723-1300
;             TELEFAX: (617) 723-8923
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2179
;             TYPE: Nucleic acid
;             STRANDEDNESS: Double
;             TOPOLOGY: Linear
;        MOLECULE TYPE: cDNA to mRNA
;        ORIGINAL SOURCE:
;             ORGANISM: Homo sapiens
;             TISSUE TYPE: Testis
;        IMMEDIATE SOURCE:
;             LIBRARY: lgt11 cDNA library, ClonTech #HL1010b
;             CLONE: pHFSHR11-11, pHFSHR15-6
;        FEATURE:
;             NAME/KEY:  protein coding region
;             LOCATION:  75 to 2159
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-474-986-1

Alignment Scores:
Pred. No.:              2.39e-06        Length:         2179
Score:                  125.00          Matches:        28
Percent Similarity:     52.83%          Conservative:   28
Best Local Similarity:  26.42%          Mismatches:     42
Query Match:            22.77%          Indels:         8
DB:                     4               Gaps:           3

US-09-930-312-2 (1-107) x US-08-474-986-1 (1-2179)

Qy      2 ValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSerIle 21
          :::|||:::|||||||||:::::::|||         ||||||   ::::::   ::::::
Db   1683 CTCAATGTCCTGGCCTTTGTGGTCATCTGTGGCTGCTATATCCACATCTACCTCACAGTG 1742

Qy     22 GlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAlaVal 41
          :::        :::   ::::::                  :::       :::
Db   1743 CGGAACCCCAACATCGTGTCCTCC-----------------TCTAGTGACACCAGGATC 1784

Qy     42 AlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPheVal 61
          |||   |||         ::::::|||:::|||   ::::|||      |||:::
Db   1785 GCCAAGCGCATGGCCATGCTCATCTTCACTGACTTCCTCTGCATGGCACCCATTTCTTTC 1844

Qy     62 ValLysIleLeuSerLeuPheArgValGlu---IleProAspThrMetThrSerTrpIle 80
            |||   :::       :::|||      |||      :::             :::
Db   1845 TTTGCCATTTCTGCCTCCCTCAAGGTGCCCCTCATCACTGTGTCCAAAGCAAAGATTCTG 1904

Qy     81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeu---Thr 99
          ::::::   |||   |||:::||||||      ||||||  ||||||   :::   |||
Db   1905 CTGGTTCTGTTTCACCCCATCAACTCCTGTGCCAACCCCTTCCTCTATGCCATCTTTACC 1964

Qy    100 ThrAsnPhePheLysAsp 105
            ||||||   :::|||
Db   1965 AAAAACTTTCGCAGAGAT 1982

RESULT 12
US-08-349-696-20
; Sequence 20, Application US/08349696
; Patent No. 5599671
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobson, Marlene A
;    APPLICANT:  Johnson, Robert G
;    APPLICANT:  Luneau, Christopher J
;    APPLICANT:  Salvatore, Christopher A
;    TITLE OF INVENTION:  Human Adenosine Receptors
;    NUMBER OF SEQUENCES:  28
```

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  P.O. Box 2000
;      CITY:  Rahway
;      STATE:  NJ
;      COUNTRY:  United States
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Macintosh IIci
;      OPERATING SYSTEM:  Macintosh
;      SOFTWARE:  Microsoft Word 5.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/349,696
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  us/08/005945
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meredith, Roy D.
;      REGISTRATION NUMBER:  30,777
;      REFERENCE/DOCKET NUMBER:  18699IA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (908)594-4678
;      TELEFAX:  (908)594-4720
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  981 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-349-696-20

Alignment Scores:
Pred. No.:              1.5e-06        Length:         981
Score:                  123.00         Matches:        29
Percent Similarity:     52.88%         Conservative:   26
Best Local Similarity:  27.88%         Mismatches:     41
Query Match:            22.40%         Indels:         8
DB:                     1              Gaps:           3

US-09-930-312-2 (1-107) x US-08-349-696-20 (1-981)

Qy       8 LeuIleIleValPheSerTyrIleThrMetPheCysSerIle----------Gln 22
           |||:::::::|||      |||:::   :::|||      |||           :::
Db     580 CTCCTCATGGTCCTCATCTACCTGGAGGTCTTCTACCTAATCCGCAAGCAGCTCAACAAG 639

Qy      23 LysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAlaValAla 42
           |||   :::      :::   :::   :::      :::|||:::|||:::   :::|||
Db     640 AAGGTGTCGGCCTCCTCCGGCGACCCGCAGAAGTACTATGGGAAGGAGCTGAAGATCGCC 699

Qy      43 AsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPheValVal 62
                     |||:::|||      |||:::   |||:::|||:::   ::::::
Db     700 AAGTCGCTGGCCCTCATCCTCTTCCTCTTTGCCCTCAGCTGGCTGCCTTTGCACATCCTC 759

Qy      63 LysIleLeuSerLeuPhe------ArgValGluIleProAspThrMetThrSerTrpIle 80
             ::::::||||||                |||      :::|||   :::|||
Db     760 AACTGCATCACCCTCTTCTGCCCGTCCTGCCACAAGCCCAGCATCCTTACC---TACATT 816

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
             ||||||           |||||||||:::||||||||||:::|||
Db     817 GCCATCTTCCTCACGCACGGCAACTCGGCCATGAACCCCATTGTCTATGCCTTCCGCATC 876

Qy     101 AsnPhePheLys 104
                |||:::
Db     877 CAGAAGTTCCGC 888

RESULT 13
US-08-233-009-20
; Sequence 20, Application US/08233009
; Patent No. 5646156
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobson, Marlene A
;    APPLICANT:  Johnson, Robert G
;    APPLICANT:  Salvatore, Christopher A
;    TITLE OF INVENTION:  INHIBITION OF EOSINOPHIL
;    TITLE OF INVENTION:  ACTIVATION THROUGH A3 ADENOSINE RECEPTOR ANTAGONISM
;    NUMBER OF SEQUENCES:  56
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  P.O.Box 2000
;      CITY:  Rahway
;      STATE:  New Jersey
;      COUNTRY:  United States
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/233,009
;      FILING DATE:  25-APR-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bencen, Gerard H
;      REGISTRATION NUMBER:  35,746
;      REFERENCE/DOCKET NUMBER:  19219
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (908) 594-3901
;      TELEFAX:  (908)594-4720
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  981 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-233-009-20

Alignment Scores:
Pred. No.:              1.5e-06        Length:         981
Score:                  123.00         Matches:        29
Percent Similarity:     52.88%         Conservative:   26
Best Local Similarity:  27.88%         Mismatches:     41
Query Match:            22.40%         Indels:         8
DB:                     1              Gaps:           3

US-09-930-312-2 (1-107) x US-08-233-009-20 (1-981)

Qy       8 LeuIleIleValPheSerTyrIleThrMetPheCysSerIle----------Gln 22
           |||:::::::|||      |||:::   :::|||      |||           :::
Db     580 CTCCTCATGGTCCTCATCTACCTGGAGGTCTTCTACCTAATCCGCAAGCAGCTCAACAAG 639

Qy      23 LysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAlaValAla 42
           |||   :::      :::   :::   :::      :::|||:::|||:::   :::|||
Db     640 AAGGTGTCGGCCTCCTCCGGCGACCCGCAGAAGTACTATGGGAAGGAGCTGAAGATCGCC 699

Qy      43 AsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPheValVal 62
                     |||:::|||      |||:::   |||:::|||:::   ::::::
Db     700 AAGTCGCTGGCCCTCATCCTCTTCCTCTTTGCCCTCAGCTGGCTGCCTTTGCACATCCTC 759

Qy      63 LysIleLeuSerLeuPhe------ArgValGluIleProAspThrMetThrSerTrpIle 80
             ::::::||||||                |||      :::|||   :::|||
Db     760 AACTGCATCACCCTCTTCTGCCCGTCCTGCCACAAGCCCAGCATCCTTACC---TACATT 816

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
             ||||||           |||||||||:::||||||||||:::|||
Db     817 GCCATCTTCCTCACGCACGGCAACTCGGCCATGAACCCCATTGTCTATGCCTTCCGCATC 876

Qy     101 AsnPhePheLys 104
                |||:::
```

```
Db    877 CAGAAGTTCCGC 888

RESULT 14
US-08-560-231-20
; Sequence 20, Application US/08560231
; Patent No. 5817760
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobson, Marlene A
;    APPLICANT:  Johnson, Robert G
;    APPLICANT:  Luneau, Christopher J
;    APPLICANT:  Salvatore, Christopher A
;    TITLE OF INVENTION:  Human Adenosine Receptors
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  P.O. Box 2000
;      CITY:  Rahway
;      STATE:  NJ
;      COUNTRY:  United States
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Macintosh IIci
;      OPERATING SYSTEM:  Macintosh
;      SOFTWARE:  Microsoft Word 5.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/560,231
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meredith, Roy D.
;      REGISTRATION NUMBER:  30,777
;      REFERENCE/DOCKET NUMBER:  18699IA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (908)594-4678
;      TELEFAX:  (908)594-4720
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  981 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-560-231-20

Alignment Scores:
Pred. No.:              1.5e-06      Length:         981
Score:                  123.00       Matches:        29
Percent Similarity:     52.88%       Conservative:   26
Best Local Similarity:  27.88%       Mismatches:     41
Query Match:            22.40%       Indels:         8
DB:                     1            Gaps:           3

US-09-930-312-2 (1-107) x US-08-560-231-20 (1-981)

Qy      8 LeuIleIleValPheSerTyrIleThrMetPheCysSerIle----------Gln 22
          |||:::::::|||      |||:::  :::|||      |||           :::
Db    580 CTCCTCATGGTCCTCATCTACCTGGAGGTCTTCTACCTAATCCGCAAGCAGCTCAACAAG 639

Qy     23 LysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAlaValAla 42
          |||   :::      :::   :::   :::      :::|||:::|||:::   :::|||
Db    640 AAGGTGTCGGCCTCCTCCGGCGACCCGCAGAAGTACTATGGGAAGGAGCTGAAGATCGCC 699

Qy     43 AsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPheValVal 62
                  |||:::|||      |||:::  |||:::|||:::   ::::::
Db    700 AAGTCGCTGGCCCTCATCCTCTTCCTCTTTGCCCTCAGCTGGCTGCCTTTGCACATCCTC 759

Qy     63 LysIleLeuSerLeuPhe------ArgValGluIleProAspThrMetThrSerTrpIle 80
            ::::::||||||                  |||      :::|||   :::|||
Db    760 AACTGCATCACCCTCTTCTGCCCGTCCTGCCACAAGCCCAGCATCCTTACC---TACATT 816

Qy     81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
          ||||||           |||||||||:::|||||||||:::|||

Db    817 GCCATCTTCCTCACGCACGGCAACTCGGCCATGAACCCCATTGTCTATGCCTTCCGCATC 876

Qy    101 AsnPhePheLys 104
                |||:::
Db    877 CAGAAGTTCCGC 888

RESULT 15
US-09-080-704A-20
; Sequence 20, Application US/09080704A
; Patent No. 6166181
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobson, Marlene A
;    APPLICANT:  Johnson, Robert G
;    APPLICANT:  Luneau, Christopher J
;    APPLICANT:  Salvatore, Christopher A
;    TITLE OF INVENTION:  Human Adenosine Receptors
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  P.O. Box 2000
;      CITY:  Rahway
;      STATE:  NJ
;      COUNTRY:  United States
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  Windows NT
;      SOFTWARE:  Word 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/080,704A
;      FILING DATE:  18 May 1998
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parr, Richard S.
;      REGISTRATION NUMBER:  32,586
;      REFERENCE/DOCKET NUMBER:  18699DB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (732)594-4958
;      TELEFAX:  (732)594-4720
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  981 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-080-704A-20

Alignment Scores:
Pred. No.:              1.5e-06      Length:         981
Score:                  123.00       Matches:        29
Percent Similarity:     52.88%       Conservative:   26
Best Local Similarity:  27.88%       Mismatches:     41
Query Match:            22.40%       Indels:         8
DB:                     4            Gaps:           3

US-09-930-312-2 (1-107) x US-09-080-704A-20 (1-981)

Qy      8 LeuIleIleValPheSerTyrIleThrMetPheCysSerIle----------Gln 22
          |||:::::::|||      |||:::  :::|||      |||           :::
Db    580 CTCCTCATGGTCCTCATCTACCTGGAGGTCTTCTACCTAATCCGCAAGCAGCTCAACAAG 639

Qy     23 LysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAlaValAla 42
          |||   :::      :::   :::   :::      :::|||:::|||:::   :::|||
Db    640 AAGGTGTCGGCCTCCTCCGGCGACCCGCAGAAGTACTATGGGAAGGAGCTGAAGATCGCC 699

Qy     43 AsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPheValVal 62
                  |||:::|||      |||:::  |||:::|||:::   ::::::
Db    700 AAGTCGCTGGCCCTCATCCTCTTCCTCTTTGCCCTCAGCTGGCTGCCTTTGCACATCCTC 759

Qy     63 LysIleLeuSerLeuPhe------ArgValGluIleProAspThrMetThrSerTrpIle 80
            ::::::||||||                  |||      :::|||   :::|||
```

```
Db     760 AACTGCATCACCCTCTTCTGCCCGTCCTGCCACAAGCCCAGCATCCTTACC---TACATT 816

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           ||||||            ||||||||||:::||||||||||:::|||
Db     817 GCCATCTTCCTCACGCACGGCAACTCGGCCATGAACCCCATTGTCTATGCCTTCCGCATC 876

Qy     101 AsnPhePheLys 104
               |||:::
Db     877 CAGAAGTTCCGC 888


Search completed: May 26, 2003, 21:21:13
Job time : 59 secs
```

GenCore version 5.1.4_p5_4578

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 26, 2003, 20:08:08 ; Search time 155 Seconds
(without alignments)
911.545 Million cell updates/sec

Title: US-09-930-312-2
Perfect score: 549
Sequence: 1 GVNLLAFLIIVFSYITMFCS.....NSALNPILYTLTTNFFKDKL 107

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09930312/runat_22052003_114615_27125/app_query.fasta_1.263
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09930312_@CGN_1_1_138_@runat_22052003_114615_27125
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	549	100.0	321	9	US-09-930-312-1	Sequence 1, Appli
2	549	100.0	2049	10	US-09-928-175-11	Sequence 11, Appl
3	549	100.0	2142	9	US-09-965-536A-5	Sequence 5, Appli
4	549	100.0	2193	10	US-09-928-175-6	Sequence 6, Appli
5	549	100.0	2214	9	US-09-965-536A-1	Sequence 1, Appli
6	549	100.0	2265	10	US-09-928-175-1	Sequence 1, Appli
7	549	100.0	2436	9	US-10-229-735-2	Sequence 2, Appli
8	549	100.0	2838	9	US-10-222-668-1	Sequence 1, Appli
9	478	87.1	2214	10	US-09-928-175-19	Sequence 19, Appl
10	478	87.1	2539	9	US-10-229-735-4	Sequence 4, Appli
11	376	68.5	530	9	US-09-989-442-33	Sequence 33, Appl
12	376	68.5	530	9	US-10-073-865-25	Sequence 25, Appl
13	376	68.5	530	9	US-10-103-313-89	Sequence 89, Appl
14	376	68.5	530	10	US-09-764-853-137	Sequence 137, App
15	376	68.5	530	10	US-09-764-877-986	Sequence 986, App
16	376	68.5	1804	10	US-09-895-686-10	Sequence 10, Appl
c 17	371	67.6	612	10	US-09-895-686-37	Sequence 37, Appl
c 18	371	67.6	1018	9	US-09-782-974C-65	Sequence 65, Appl
19	234	42.6	1015	9	US-10-270-333-20	Sequence 20, Appl
20	225	41.0	1080	9	US-10-270-333-26	Sequence 26, Appl
21	225	41.0	3187	9	US-10-270-333-25	Sequence 25, Appl
22	181	33.0	3574	9	US-10-270-333-19	Sequence 19, Appl
c 23	149.5	27.2	432	10	US-09-895-686-62	Sequence 62, Appl
24	142	25.9	2586	10	US-09-804-551B-19	Sequence 19, Appl
25	138.5	25.2	2100	10	US-09-804-626-5	Sequence 5, Appli
26	136.5	24.9	2902	10	US-09-877-804-1	Sequence 1, Appli
27	125	22.8	2088	10	US-09-804-626-7	Sequence 7, Appli
28	123	22.4	1942	9	US-09-800-274-1	Sequence 1, Appli
29	123	22.4	3349	9	US-10-270-336-4	Sequence 4, Appli
30	122	22.2	2366	10	US-09-877-804-5	Sequence 5, Appli
31	120	21.9	2487	9	US-10-270-336-1	Sequence 1, Appli
32	120	21.9	3438	9	US-10-176-847-89	Sequence 89, Appl
33	120	21.9	119596	9	US-10-270-336-3	Sequence 3, Appli
34	119.5	21.8	1739	9	US-09-800-274-3	Sequence 3, Appli
35	118.5	21.6	2512	9	US-10-045-624B-1	Sequence 1, Appli
36	113	20.6	572	9	US-09-925-299-296	Sequence 296, App
37	113	20.6	572	10	US-09-925-299-296	Sequence 296, App
38	109	19.9	2703	9	US-09-984-271-75	Sequence 75, Appl
39	109	19.9	2714	9	US-09-925-299-202	Sequence 202, App
40	109	19.9	2714	10	US-09-925-299-202	Sequence 202, App
41	109	19.9	3429	9	US-09-822-846-124	Sequence 124, App
42	109	19.9	3688	10	US-09-862-767A-1	Sequence 1, Appli
43	105	19.1	422	9	US-09-918-995-34626	Sequence 34626, A
44	105	19.1	1032	9	US-09-971-228-2	Sequence 2, Appli
45	105	19.1	1089	10	US-09-993-844-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-09-930-312-1
; Sequence 1, Application US/09930312
; Publication No. US20030032019A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn Company
; APPLICANT: Lind, Peter
; TITLE OF INVENTION: No. US20030032019Alel G Protein-Coupled Receptors
; FILE REFERENCE: 00329.PCT1
; CURRENT APPLICATION NUMBER: US/09/930,312
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: USSN 60/225,262
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 321
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-312-1

Alignment Scores:

Pred. No.:	2.89e-61	Length:	321
Score:	549.00	Matches:	107
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

```
US-09-930-312-2 (1-107) x US-09-930-312-1 (1-321)

Qy      1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 60

Qy     21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 120

Qy     41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 180

Qy     61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 240

Qy     81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 300

Qy    101 AsnPhePheLysAspLysLeu 107
          |||||||||||||||||||||
Db    301 AACTTTTTTAAGGACAAGTTG 321

RESULT 2
US-09-928-175-11
; Sequence 11, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
;  APPLICANT: Paszty, Christopher J.
;  APPLICANT:  Gong, Jianhua
;  APPLICANT:  Daugherty, Betsy
;  APPLICANT:  Rogers, No. US20020123618A1ma
;  TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
;  TITLE OF INVENTION:  Uses Thereof
;  FILE REFERENCE: 00-1229
;  CURRENT APPLICATION NUMBER: US/09/928,175
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: 60/224,455
;  PRIOR FILING DATE: 2000-08-10
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 2049
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(2046)
;   NAME/KEY: sig_peptide
;   LOCATION: (1)..(108)
US-09-928-175-11

Alignment Scores:
Pred. No.:              4e-60          Length:         2049
Score:                  549.00         Matches:        107
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            100.00%        Indels:         0
DB:                     10             Gaps:           0

US-09-930-312-2 (1-107) x US-09-928-175-11 (1-2049)

Qy      1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1570 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1629

Qy     21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1630 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1689

Qy     41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1690 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1749

Qy     61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1750 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1809

Qy     81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1810 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 1869

Qy    101 AsnPhePheLysAspLysLeu 107
          |||||||||||||||||||||
Db   1870 AACTTTTTTAAGGACAAGTTG 1890

RESULT 3
US-09-965-536A-5
; Sequence 5, Application US/09965536A
; Publication No. US20030027323A1
; GENERAL INFORMATION:
;  APPLICANT: FEDER, J. N.
;  APPLICANT:  MINTIER, G.
;  APPLICANT:  RAMANATHAN, C. S.
;  APPLICANT:  HAWKEN, D. R.
;  TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMY5,
;  TITLE OF INVENTION:  EXPRESSED HIGHLY IN BRAIN AND OVARIAN TISSUES
;  FILE REFERENCE: DO041NP
;  CURRENT APPLICATION NUMBER: US/09/965,536A
;  CURRENT FILING DATE:  2001-09-26
;  PRIOR APPLICATION NUMBER: 60/235,713
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: 60/261,781
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: 60/306,605
;  PRIOR FILING DATE: 2001-07-19
;  PRIOR APPLICATION NUMBER: 60/310,436
;  PRIOR FILING DATE: 2001-08-03
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 2142
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-965-536A-5

Alignment Scores:
Pred. No.:              4.26e-60       Length:         2142
Score:                  549.00         Matches:        107
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            100.00%        Indels:         0
DB:                     9              Gaps:           0

US-09-930-312-2 (1-107) x US-09-965-536A-5 (1-2142)

Qy      1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1663 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1722

Qy     21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1723 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1782

Qy     41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1783 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1842

Qy     61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1843 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1902
```

```
Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1903 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 1962

Qy     101 AsnPhePheLysAspLysLeu 107
           |||||||||||||||||||
Db    1963 AACTTTTTTAAGGACAAGTTG 1983

RESULT 4
US-09-928-175-6
; Sequence 6, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
;  APPLICANT: Paszty, Christopher J.
;  APPLICANT:  Gong, Jianhua
;  APPLICANT:  Daugherty, Betsy
;  APPLICANT:  Rogers, No. US20020123618A1ma
;  TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
;  TITLE OF INVENTION:  Uses Thereof
;  FILE REFERENCE: 00-1229
;  CURRENT APPLICATION NUMBER: US/09/928,175
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: 60/224,455
;  PRIOR FILING DATE: 2000-08-10
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 2193
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(2190)
;   NAME/KEY: sig_peptide
;   LOCATION: (1)..(108)
US-09-928-175-6

Alignment Scores:
Pred. No.:              4.41e-60      Length:         2193
Score:                  549.00        Matches:        107
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            100.00%       Indels:         0
DB:                     10            Gaps:           0

US-09-930-312-2 (1-107) x US-09-928-175-6 (1-2193)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1714 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1773

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1774 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1833

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1834 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1893

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1894 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1953

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1954 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2013

Qy     101 AsnPhePheLysAspLysLeu 107
           |||||||||||||||||||
Db    2014 AACTTTTTTAAGGACAAGTTG 2034

RESULT 5
US-09-965-536A-1
; Sequence 1, Application US/09965536A
; Publication No. US20030027323A1
; GENERAL INFORMATION:
;  APPLICANT: FEDER, J. N.
;  APPLICANT:  MINTIER, G.
;  APPLICANT:  RAMANATHAN, C. S.
;  APPLICANT:  HAWKEN, D. R.
;  TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMY5,
;  TITLE OF INVENTION:  EXPRESSED HIGHLY IN BRAIN AND OVARIAN TISSUES
;  FILE REFERENCE: DO041NP
;  CURRENT APPLICATION NUMBER: US/09/965,536A
;  CURRENT FILING DATE:  2001-09-26
;  PRIOR APPLICATION NUMBER: 60/235,713
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: 60/261,781
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: 60/306,605
;  PRIOR FILING DATE: 2001-07-19
;  PRIOR APPLICATION NUMBER: 60/310,436
;  PRIOR FILING DATE: 2001-08-03
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 2214
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-965-536A-1

Alignment Scores:
Pred. No.:              4.47e-60      Length:         2214
Score:                  549.00        Matches:        107
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            100.00%       Indels:         0
DB:                     9             Gaps:           0

US-09-930-312-2 (1-107) x US-09-965-536A-1 (1-2214)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1735 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1794

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1795 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1854

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1855 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1914

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1915 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 1974

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1975 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2034

Qy     101 AsnPhePheLysAspLysLeu 107
           |||||||||||||||||||
Db    2035 AACTTTTTTAAGGACAAGTTG 2055

RESULT 6
US-09-928-175-1
; Sequence 1, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
;  APPLICANT: Paszty, Christopher J.
;  APPLICANT:  Gong, Jianhua
;  APPLICANT:  Daugherty, Betsy
;  APPLICANT:  Rogers, No. US20020123618A1ma
;  TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
;  TITLE OF INVENTION:  Uses Thereof
```

```
;  FILE REFERENCE: 00-1229
;  CURRENT APPLICATION NUMBER: US/09/928,175
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: 60/224,455
;  PRIOR FILING DATE: 2000-08-10
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 2265
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(2262)
;   NAME/KEY: sig_peptide
;   LOCATION: (1)..(108)
US-09-928-175-1

Alignment Scores:
Pred. No.:              4.61e-60        Length:         2265
Score:                  549.00          Matches:        107
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     10              Gaps:           0

US-09-930-312-2 (1-107) x US-09-928-175-1 (1-2265)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1786 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1845

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1846 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1905

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1906 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 1965

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1966 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 2025

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2026 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2085

Qy     101 AsnPhePheLysAspLysLeu 107
           |||||||||||||||||||||
Db    2086 AACTTTTTTAAGGACAAGTTG 2106

RESULT 7
US-10-229-735-2
; Sequence 2, Application US/10229735
; Publication No. US20030082650A1
; GENERAL INFORMATION:
;  APPLICANT: Baylor College of Medicine
;  APPLICANT:  Agoulnik, Alexander I.
;  TITLE OF INVENTION: The GREAT Gene and Protein
;  FILE REFERENCE: 7572/73263
;  CURRENT APPLICATION NUMBER: US/10/229,735
;  CURRENT FILING DATE:  2002-08-29
;  PRIOR APPLICATION NUMBER: 60/315,696
;  PRIOR FILING DATE: 2001-08-30
;  PRIOR APPLICATION NUMBER: 60/351,432
;  PRIOR FILING DATE: 2002-01-28
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 2436
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-229-735-2

Alignment Scores:
Pred. No.:              5.11e-60        Length:         2436
Score:                  549.00          Matches:        107
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-930-312-2 (1-107) x US-10-229-735-2 (1-2436)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1856 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1915

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1916 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 1975

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1976 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 2035

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2036 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 2095

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2096 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2155

Qy     101 AsnPhePheLysAspLysLeu 107
           |||||||||||||||||||||
Db    2156 AACTTTTTTAAGGACAAGTTG 2176

RESULT 8
US-10-222-668-1
; Sequence 1, Application US/10222668
; Publication No. US20030088884A1
; GENERAL INFORMATION:
;  APPLICANT: Hsu, Sheau Yu
;  APPLICANT:  Hsueh, Aaron
;  TITLE OF INVENTION: Mammalian Relaxin Receptor
;  FILE REFERENCE: STAN-239 WO
;  CURRENT APPLICATION NUMBER: US/10/222,668
;  CURRENT FILING DATE:  2002-08-15
;  PRIOR APPLICATION NUMBER: 60/313,259
;  PRIOR FILING DATE: 2002-08-17
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 2838
;   TYPE: DNA
;   ORGANISM: H. sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (107)...(2369)
US-10-222-668-1

Alignment Scores:
Pred. No.:              6.35e-60        Length:         2838
Score:                  549.00          Matches:        107
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-930-312-2 (1-107) x US-10-222-668-1 (1-2838)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1892 GGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCCTATATTACTATGTTCTGTTCC 1951
```

```
Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1952 ATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAATTGTTTTGGAAGAGAGGTGGCT 2011

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2012 GTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCCATCTGCTGGATTCCTGTATTT 2071

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2072 GTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCAGACACAATGACTTCCTGGATA 2131

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2132 GTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCAATCCTCTATACTCTCACAACC 2191

Qy     101 AsnPhePheLysAspLysLeu 107
           |||||||||||||||||||
Db    2192 AACTTTTTTAAGGACAAGTTG 2212

RESULT 9
US-09-928-175-19
; Sequence 19, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
;  APPLICANT: Paszty, Christopher J.
;  APPLICANT:  Gong, Jianhua
;  APPLICANT:  Daugherty, Betsy
;  APPLICANT:  Rogers, No. US20020123618Alma
;  TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
;  TITLE OF INVENTION:  Uses Thereof
;  FILE REFERENCE: 00-1229
;  CURRENT APPLICATION NUMBER: US/09/928,175
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: 60/224,455
;  PRIOR FILING DATE: 2000-08-10
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
;   LENGTH: 2214
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(2211)
;   NAME/KEY: sig_peptide
;   LOCATION: (1)..(57)
US-09-928-175-19

Alignment Scores:
Pred. No.:              4.95e-51      Length:          2214
Score:                  478.00        Matches:         92
Percent Similarity:     93.46%        Conservative:    8
Best Local Similarity:  85.98%        Mismatches:      7
Query Match:            87.07%        Indels:          0
DB:                     10            Gaps:            0

US-09-930-312-2 (1-107) x US-09-928-175-19 (1-2214)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           ||||||||||||||||||||||||:::|||||||   |||||||:::|||||||||||||||
Db    1735 GGTGTGAACTTGCTGGCTTTCCTCGTCATCGTGATTTCCTATGTCACCATGTTCTGCTCC 1794

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           |||   ||||||||||||||||||   |||||||:::      |||:::||||||||||
Db    1795 ATTCATAAAACAGCCCTTCAGACTGCAGAAGTGAGGAGCCACATCGGGAAGGAGGTGGCT 1854

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1855 GTTGCAAACCGGTTCTTTTTTATCGTGTTCTCTGATGCCATCTGCTGGATCCCTGTGTTT 1914

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           ||||||||||||||||||||   :::|||||||||||   |||:::||||||||||||||
Db    1915 GTCGTTAAGATCCTGTCTCTCCTTCAAGTGGAGATACCAGGCACAATCACTTCCTGGATC 1974

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           |||:::|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1975 GTGGTTTTTTTCCTTCCGGTGAACAGCGCCTTAAACCCCATCCTCTACACTCTGACGACC 2034

Qy     101 AsnPhePheLysAspLysLeu 107
           :::|||||||||||||||||||
Db    2035 TCCTTTTTTAAGGACAAGTTG 2055

RESULT 10
US-10-229-735-4
; Sequence 4, Application US/10229735
; Publication No. US20030082650A1
; GENERAL INFORMATION:
;  APPLICANT: Baylor College of Medicine
;  APPLICANT:  Agoulnik, Alexander I.
;  TITLE OF INVENTION: The GREAT Gene and Protein
;  FILE REFERENCE: 7572/73263
;  CURRENT APPLICATION NUMBER: US/10/229,735
;  CURRENT FILING DATE:  2002-08-29
;  PRIOR APPLICATION NUMBER: 60/315,696
;  PRIOR FILING DATE: 2001-08-30
;  PRIOR APPLICATION NUMBER: 60/351,432
;  PRIOR FILING DATE: 2002-01-28
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 2539
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-229-735-4

Alignment Scores:
Pred. No.:              6.01e-51      Length:          2539
Score:                  478.00        Matches:         92
Percent Similarity:     93.46%        Conservative:    8
Best Local Similarity:  85.98%        Mismatches:      7
Query Match:            87.07%        Indels:          0
DB:                     9             Gaps:            0

US-09-930-312-2 (1-107) x US-10-229-735-4 (1-2539)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           ||||||||||||||||||||||||:::|||||||   |||||||:::|||||||||||||||
Db    1928 GGTGTGAACTTGCTGGCTTTCCTCGTCATCGTGATTTCCTATGTCACCATGTTCTGCTCC 1987

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           |||   ||||||||||||||||||   |||||||:::      |||:::||||||||||
Db    1988 ATTCATAAAACAGCCCTTCAGACTGCAGAAGTGAGGAGCCACATCGGGAAGGAGGTGGCT 2047

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2048 GTTGCAAACCGGTTCTTTTTTATCGTGTTCTCTGATGCCATCTGCTGGATCCCTGTGTTT 2107

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           ||||||||||||||||||||   :::|||||||||||   |||:::||||||||||||||
Db    2108 GTCGTTAAGATCCTGTCTCTCCTTCAAGTGGAGATACCAGGCACAATCACTTCCTGGATC 2167

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           |||:::|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2168 GTGGTTTTTTTCCTTCCGGTGAACAGCGCCTTAAACCCCATCCTCTACACTCTGACGACC 2227

Qy     101 AsnPhePheLysAspLysLeu 107
           :::|||||||||||||||||||
Db    2228 TCCTTTTTTAAGGACAAGTTG 2248

RESULT 11
US-09-989-442-33
; Sequence 33, Application US/09989442
; Publication No. US20030013649A1
; GENERAL INFORMATION:
```

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJZ08
; CURRENT APPLICATION NUMBER: US/09/989,442
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245

```
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,244
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,217
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,211
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,215
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,264
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,214
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/249,297
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/232,400
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/231,242
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/232,081
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/232,080
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/231,414
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/231,244
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/233,064
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/233,063
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/232,397
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/232,399
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/232,401
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: 60/241,808
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/241,826
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/241,786
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/241,221
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/246,475
;  PRIOR FILING DATE: 2000-11-08
;  PRIOR APPLICATION NUMBER: 60/231,243
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/233,065
;  PRIOR FILING DATE: 2000-09-14

Alignment Scores:
Pred. No.:                6.42e-39      Length:          530
Score:                    376.00        Matches:         69
Percent Similarity:       83.18%        Conservative:    20
Best Local Similarity:    64.49%        Mismatches:      18
Query Match:              68.49%        Indels:          0
DB:                       9             Gaps:            0

US-09-930-312-2 (1-107) x US-09-989-442-33 (1-530)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           |||:::||||||   ||||||:::||||||||||||||||||   :::||||||   |||
Db      32 GGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGT 91

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           :::   ::::::|||:::     ||||||:::||||||         :::|||:::
Db      92 GTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATC 151

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           :::|||   ||||||||||||||||||||:::||||||:::||||||||||||||:::|||
Db     152 CTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTT 211

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           |||||||||   |||||||||   :::||||||||||||   |||:::||||||||:::
Db     212 GTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTA 271

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           |||||||||   ||||||:::||||||||||||||||||||||||||||||||||||||||
Db     272 GTGATTTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACA 331

Qy     101 AsnPhePheLysAspLysLeu 107
                 |||||||:::   :::
Db     332 AGACCATTTAAAGAAATGATT 352

RESULT 12
US-10-073-865-25
; Sequence 25, Application US/10073865
; Publication No. US20030044904A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PJZ09C1
;  CURRENT APPLICATION NUMBER: US/10/073,865
;  CURRENT FILING DATE:  2002-02-14
;  Prior Application removed - See file Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 154
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
;   LENGTH: 530
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-073-865-25

Alignment Scores:
Pred. No.:                6.42e-39      Length:          530
Score:                    376.00        Matches:         69
Percent Similarity:       83.18%        Conservative:    20
Best Local Similarity:    64.49%        Mismatches:      18
Query Match:              68.49%        Indels:          0
DB:                       9             Gaps:            0

US-09-930-312-2 (1-107) x US-10-073-865-25 (1-530)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           |||:::||||||   ||||||:::||||||||||||||||||   :::||||||   |||
Db      32 GGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGT 91

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           :::   ::::::|||:::     ||||||:::||||||         :::|||:::
Db      92 GTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATC 151

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           :::|||   ||||||||||||||||||||:::||||||:::||||||||||||||:::|||
Db     152 CTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTT 211

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           |||||||||   |||||||||   :::||||||||||||   |||:::||||||||:::
Db     212 GTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTA 271

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           |||||||||   ||||||:::||||||||||||||||||||||||||||||||||||||||
Db     272 GTGATTTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACA 331

Qy     101 AsnPhePheLysAspLysLeu 107
                 |||||||:::   :::
Db     332 AGACCATTTAAAGAAATGATT 352

RESULT 13
US-10-103-313-89
; Sequence 89, Application US/10103313
; Publication No. US20030082758A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
```

```
;  FILE REFERENCE: PJZ07C1
;  CURRENT APPLICATION NUMBER: US/10/103,313
;  CURRENT FILING DATE:  2002-03-12
;  NUMBER OF SEQ ID NOS: 653
;       Prior Application removed - See File Wrapper or Palm
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 89
;   LENGTH: 530
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-103-313-89
```

Alignment Scores:

Pred. No.:	6.42e-39	Length:	530
Score:	376.00	Matches:	69
Percent Similarity:	83.18%	Conservative:	20
Best Local Similarity:	64.49%	Mismatches:	18
Query Match:	68.49%	Indels:	0
DB:	9	Gaps:	0

US-09-930-312-2 (1-107) x US-10-103-313-89 (1-530)

```
Qy      1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
          |||:::||||||   ||||||:::||||||||||||||||||   :::||||||   |||
Db     32 GGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGT 91

Qy     21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
          :::   ::::::|||:::      ||||||:::||||||         :::|||:::
Db     92 GTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATC 151

Qy     41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
          :::|||   ||||||||||||||||||||:::||||||:::||||||||||||:::|||
Db    152 CTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTT 211

Qy     61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
          |||||||||   |||||||||   :::||||||||||||   |||:::||||||||:::
Db    212 GTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTA 271

Qy     81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
          |||||||||   ||||||:::||||||||||||||||||||||||||||||||||||||
Db    272 GTGATTTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACA 331

Qy    101 AsnPhePheLysAspLysLeu 107
                |||||||:::   :::
Db    332 AGACCATTTAAAGAAATGATT 352
```

RESULT 14

```
US-09-764-853-137
; Sequence 137, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PJZ06
;  CURRENT APPLICATION NUMBER: US/09/764,853
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 939
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 137
;   LENGTH: 530
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-853-137
```

Alignment Scores:

Pred. No.:	6.42e-39	Length:	530
Score:	376.00	Matches:	69
Percent Similarity:	83.18%	Conservative:	20
Best Local Similarity:	64.49%	Mismatches:	18
Query Match:	68.49%	Indels:	0
DB:	10	Gaps:	0

US-09-930-312-2 (1-107) x US-09-764-853-137 (1-530)

```
Qy      1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
          |||:::||||||   ||||||:::||||||||||||||||||   :::||||||   |||
Db     32 GGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGT 91

Qy     21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
          :::   ::::::|||:::      ||||||:::||||||         :::|||:::
Db     92 GTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATC 151

Qy     41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
          :::|||   ||||||||||||||||||||:::||||||:::||||||||||||:::|||
Db    152 CTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTT 211

Qy     61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
          |||||||||   |||||||||   :::||||||||||||   |||:::||||||||:::
Db    212 GTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTA 271

Qy     81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
          |||||||||   ||||||:::||||||||||||||||||||||||||||||||||||||
```

RESULT 15

```
US-09-764-877-986
; Sequence 986, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC005
;  CURRENT APPLICATION NUMBER: US/09/764,877
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 4031
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 986
;   LENGTH: 530
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-877-986
```

Alignment Scores:

Pred. No.:	6.42e-39	Length:	530
Score:	376.00	Matches:	69
Percent Similarity:	83.18%	Conservative:	20
Best Local Similarity:	64.49%	Mismatches:	18
Query Match:	68.49%	Indels:	0
DB:	10	Gaps:	0

US-09-930-312-2 (1-107) x US-09-764-877-986 (1-530)

```
Qy      1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
          |||:::||||||   ||||||:::||||||||||||||||||   :::||||||   |||
Db     32 GGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGT 91

Qy     21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
          :::   ::::::|||:::      ||||||:::||||||         :::|||:::
Db     92 GTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATC 151

Qy     41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
          :::|||   ||||||||||||||||||||:::||||||:::||||||||||||:::|||
Db    152 CTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTT 211

Qy     61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
          |||||||||   |||||||||   :::||||||||||||   |||:::||||||||:::
Db    212 GTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTA 271

Qy     81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
          |||||||||   ||||||:::||||||||||||||||||||||||||||||||||||||
```

```
Db    272 GTGATTTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACA 331

Qy    101 AsnPhePheLysAspLysLeu 107
              ||||||:::   :::
Db    332 AGACCATTTAAAGAAATGATT 352


Search completed: May 26, 2003, 21:24:03
Job time : 162 secs
```

GenCore version 5.1.4_p5_4578

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 26, 2003, 20:03:57 ; Search time 1716 Seconds
(without alignments)
1009.858 Million cell updates/sec

Title: US-09-930-312-2
Perfect score: 549
Sequence: 1 GVNLLAFLIIVFSYITMFCS.....NSALNPILYTLTTNFFKDKL 107

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09930312/runat_22052003_114614_27077/app_query.fasta_1.263
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09930312_@CGN_1_1_1906_@runat_22052003_114614_27077 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	375	68.3	744	12	BG304121	BG304121 fl30f05.x
2	361	65.8	841	14	BQ228832	BQ228832 AGENCOURT
3	316	57.6	197	12	BF942735	BF942735 EST-CD15N
4	288	52.5	235	10	AW436170	AW436170 75600 MAR
5	162	29.5	456	12	BF021857	BF021857 uy58h08.y
6	162	29.5	574	10	BB632910	BB632910 BB632910
7	142	25.9	573	13	BI630907	BI630907 RH60250.5
8	142	25.9	610	13	BI628622	BI628622 RH57122.5
9	137.5	25.0	511	14	BM960298	BM960298 cihA4L14S
10	137.5	25.0	522	10	AV995043	AV995043 AV995043
c 11	137.5	25.0	585	10	AV883650	AV883650 AV883650
12	132.5	24.1	676	10	BB553437	BB553437 BB553437
13	132.5	24.1	1113	13	BM454054	BM454054 AGENCOURT
14	132.5	24.1	1142	13	BM461149	BM461149 AGENCOURT
15	127.5	23.2	2372	11	AK013534	AK013534 Mus muscu
16	126	23.0	568	13	BI608244	BI608244 RH12435.5
17	126	23.0	2360	11	AK016635	AK016635 Mus muscu
18	125	22.8	548	12	BF075894	BF075894 225322 MA
19	124	22.6	935	17	CNS01M4J	AL150468 Anopheles
20	123	22.4	896	14	BQ225080	BQ225080 AGENCOURT
21	123	22.4	973	12	BG423878	BG423878 602449840
22	123	22.4	975	14	BQ958975	BQ958975 AGENCOURT
23	123	22.4	981	12	BG168804	BG168804 602320111
24	123	22.4	982	12	BF159363	BF159363 601770339
c 25	122	22.2	563	14	BM962037	BM962037 cihA4L14T
c 26	120.5	21.9	613	10	BE179105	BE179105 RC0-HT061
c 27	120.5	21.9	1101	17	CNS0181F	AL108765 Drosophil
c 28	120	21.9	655	12	BE985448	BE985448 UI-M-CG0p
29	118.5	21.6	3015	11	AK014519	AK014519 Mus muscu
30	117	21.3	518	10	AW660289	AW660289 98243 MAR
c 31	116.5	21.2	1050	17	CNS04V9T	AL308810 Tetraodon
32	115	20.9	592	14	BM783394	BM783394 K-EST0061
c 33	115	20.9	669	10	AW367357	AW367357 MR0-HT016
34	115	20.9	704	14	BM722841	BM722841 UI-E-EJ0-
35	115	20.9	786	12	BG202798	BG202798 RST22165
c 36	115	20.9	985	9	AL571416	AL571416 AL571416
c 37	114.5	20.9	885	17	CNS030UT	AL222734 Tetraodon
38	113.5	20.7	528	14	BQ667379	BQ667379 pb60e12.y
39	113	20.6	510	12	BF516393	BF516393 UI-H-BW1-
40	112	20.4	571	14	BM741294	BM741294 K-EST0013
41	112	20.4	589	14	BM751805	BM751805 K-EST0028
42	112	20.4	928	9	AL545857	AL545857 AL545857
c 43	112	20.4	944	9	AL576349	AL576349 AL576349
c 44	112	20.4	1002	9	AL576584	AL576584 AL576584
45	111.5	20.3	538	13	BM130064	BM130064 pb26a08.y

ALIGNMENTS

RESULT 1
BG304121/c
LOCUS BG304121 744 bp mRNA linear EST 23-FEB-2001
DEFINITION fl30f05.x1 Sugano Kawakami zebrafish DRA Danio rerio cDNA clone 3815552 3' similar to TR:Q9VBP0 Q9VBP0 CG5042 PROTEIN. ;, mRNA sequence.
ACCESSION BG304121
VERSION BG304121.1 GI:13101648
KEYWORDS EST.
SOURCE zebrafish.
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes ; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 744)

```
  AUTHORS   Sugano,S., Kawakami,K., Johnson,S., Li,F., Marra,M., Eddy,S.,
            Hillier,L., Clifton,S., Allen,M., Gibbons,M., Jost,S., Kucaba,T.,
            Martin,J., Pape,D., Steptoe,M., Underwood,K., Theising,B., Ritter
            ,E., Bowers,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: S.L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA
            Sequencing by: Washington University Genome Sequencing Center Clone
            distribution information can be found through the I.M.A.G.E.
            Consortium/LLNL, send email to: info@image.llnl.gov
            Seq primer: T7 from Gibco
            High quality sequence stop: 510.
FEATURES             Location/Qualifiers
     source          1. .744
                     /organism="Danio rerio"
                     /strain="AB"
                     /db_xref="taxon:7955"
                     /clone="3815552"
                     /clone_lib="Sugano Kawakami zebrafish DRA"
                     /sex="mixed (one male and one female, including
                     unfertilized eggs)"
                     /dev_stage="adult"
                     /lab_host="DH10B (phage resistant)"
                     /note="Vector: pME18S-FL3; Site_1: DraIII (CACTGTGTG);
                     Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed
                     with an oligo(dT) primer  [ATGTGGCCTTTTTTTTTTTTTTTT};
                     double-stranded cDNA was ligated to a  DraIII adaptor
                     [TGTTGGCCTACTGG], digested and cloned into distinct DraIII
                     sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site
                     CACCATGTG). XhoI should be used to isolate the cDNA
                     insert.  Size selection was performed to exclude fragments
                     <1.5kb.  Library constructed by Dr. Sumio Sugano
                     (University of Tokyo Institute of Medical Science) and
                     kindly donated by Dr. Koichi Kawakami.  Custom primers for
                     sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end
                     primer CGACCTGCAGCTCGAGCACA.  "
BASE COUNT      204 a    158 c    186 g    196 t
ORIGIN

Alignment Scores:
Pred. No.:              3.8e-34       Length:         744
Score:                  375.00        Matches:        71
Percent Similarity:     83.18%        Conservative:   18
Best Local Similarity:  66.36%        Mismatches:     18
Query Match:            68.31%        Indels:         0
DB:                     12            Gaps:           0

US-09-930-312-2 (1-107) x BG304121 (1-744)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           |||:::      ||||||||||||||||||||||   ||||||   :::||||||   |||
Db     738 GGACTTACCCGGCTGGCTTTCCTCATCATCGTGGTCTCGTACTCCAGCATGTTTTACTCC 679

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           |||   ||||||   ::::::   |||||:::               |||:::||||||
Db     678 ATCTATAAGACGGGCATCAGAGCCACAGAGCTGCCGGGTCGACTGCACAGAGACGTGGCC 619

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           :::|||:::|||||||||||||||||||||||||||||   ||||||||||||:::|||
Db     618 ATGGCACACCGCTTCTTCTTCATCGTGTTTTCTGATGCCTTCTGCTGGATTCCCATATTC 559

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           :::||||||||||||||||||      ||||||||||||   |||:::||||||||:::
Db     558 ATGGTCAAAATCCTCTCTCTAATGGAGGTGGAGATACCAGGAACCATCACATCTTGGGTG 499

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           |||||||||   ||||||:::||||||||||||||||||||||||||||||||||||||||

Db     498 GTGATCTTCATCCTGCCCATCAACAGTGCCTTAAACCCCATCCTTTACACACTGACCACC 439

Qy     101 AsnPhePheLysAspLysLeu 107
           :::||||||:::::::::::::
Db     438 AGTTTTTTCAGAGAGCAGGTG 418

RESULT 2
BQ228832
LOCUS       BQ228832                841 bp    mRNA    linear   EST 02-MAY-2002
DEFINITION  AGENCOURT_7522585 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6051251
            5', mRNA sequence.
ACCESSION   BQ228832
VERSION     BQ228832.1  GI:20410232
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 841)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13304  row: m  column: 12
            High quality sequence start: 87
            High quality sequence stop: 603.
FEATURES             Location/Qualifiers
     source          1. .841
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6051251"
                     /clone_lib="NIH_MGC_72"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
BASE COUNT      240 a    160 c    170 g    269 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:              1.89e-32      Length:         841
Score:                  361.00        Matches:        68
Percent Similarity:     82.41%        Conservative:   21
Best Local Similarity:  62.96%        Mismatches:     18
Query Match:            65.76%        Indels:         1
DB:                     14            Gaps:           0

US-09-930-312-2 (1-107) x BQ228832 (1-841)

Qy       1 GlyValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSer 20
           |||:::||||||   ||||||:::||||||||||||||||||   :::||||||   |||
Db     335 GGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGT 394

Qy      21 IleGlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           :::   ::::::|||:::      ||||||:::||||||          :::|||:::
Db     395 GTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATC 454

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           :::|||   ||||||||||||||||||||:::|||||:::||||||||||||:::|||
Db     455 CTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTT 514

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
           |||||||||   |||||||||   :::||||||||||||   |||:::||||||||:::
```

```
Db    515 GTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTA 574

Qy     81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeu-ThrTh 100
          |||||||||   ||||||:::|||||||||||||||||||||||||||||||| |||
Db    575 GTGATTTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGAACCCA 634

Qy    100 rAsnPhePheLysAspLysLeu 107
          :::   ||||||:::   :::
Db    635 AGACCATTTTAAAGAAATGATT 656

RESULT 3
BF942735
LOCUS       BF942735                197 bp    mRNA    linear   EST 15-SEP-2001
DEFINITION  EST-CD15N-040 human CD15+ myeloid progenitor cells cDNA Library
            Homo sapiens cDNA 3', mRNA sequence.
ACCESSION   BF942735
VERSION     BF942735.1  GI:15624134
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 197)
  AUTHORS   Chen,J., Lee,S., Zhou,G., Rowley,J.D. and Wang,S.M.
  TITLE     A high-throughput GLGI procedure for converting large number of
            SAGE tag sequences into 3' ESTs
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Wang SM
            Hem/Onc
            University of Chicago Medical Center
            5841 S. Maryland Ave., MC2115, Chicago, IL 60637, USA
            Tel: 773-702-6788
            Fax: 773-702-3002
            Email: swang1@midway.uchicago.edu
            This EST fragment was amplified from human CD15+ myeloid progenitor
            cells cDNA Library with GLGI technique (Generation of Longer cDNA
            fragments from SAGE tags for Gene Identification, Proc. Natl. Acad.
            Sci. USA 97, 349, 2000), which starts from the 3' end till the last
            CATG site of the target cDNA sequence.
            Seq primer: M13 Forward.
FEATURES             Location/Qualifiers
     source          1. .197
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="human CD15+ myeloid progenitor cells cDNA
                     Library"
                     /tissue_type="bone marrow"
                     /cell_type="CD15+ myeloid progenitor cells"
                     /note="Organ: thyroid; Vector: pAMP10; mRNA made from
                     thyroid carcinoma, cDNA made by oligo-dT  priming.
                     Non-directionally cloned into UDG sites. Size-selected on
                     agarose gel, average insert size 500 bp. Primary library.
                     cDNA Library Preparation: David B. Krizman, Ph.D.
                     REFERENCE: Krizman et al. (1996) Cancer Research
                     56:5380-5383."
BASE COUNT       44 a     38 c     37 g     78 t
ORIGIN

Alignment Scores:
Pred. No.:              8.81e-28        Length:         197
Score:                  316.00          Matches:        61
Percent Similarity:     98.39%          Conservative:   0
Best Local Similarity:  98.39%          Mismatches:     1
Query Match:            57.56%          Indels:         0
DB:                     12              Gaps:           0

US-09-930-312-2 (1-107) x BF942735 (1-197)

Qy     12 PheSerTyrIleThrMetPheCysSerIleGlnLysThrAlaLeuGlnThrThrGluVal 31
          |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     10 TTTCCCTATATTACTATGTTCTGTTCCATTCAAAAAACCGCCTTGCAGACCACAGAAGTA 69

Qy     32 ArgAsnCysPheGlyArgGluValAlaValAlaAsnArgPhePhePheIleValPheSer 51
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     70 AGGAATTGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTTATAGTGTTCTCT 129

Qy     52 AspAlaIleCysTrpIleProValPheValValLysIleLeuSerLeuPheArgValGlu 71
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    130 GATGCCATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAA 189

Qy     72 IlePro 73
          ||||||
Db    190 ATACCA 195

RESULT 4
AW436170
LOCUS       AW436170                235 bp    mRNA    linear   EST 09-JUL-2000
DEFINITION  75600 MARC 2PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION   AW436170
VERSION     AW436170.1  GI:6971476
KEYWORDS    EST.
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 235)
  AUTHORS   Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
            Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
            and Keele,J.W.
  TITLE     Design and use of two pooled tissue normalized cDNA libraries for
            EST discovery in swine
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 20
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 34  row: F  column: 22
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .235
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
                     /clone_lib="MARC 2PIG"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
                     Library made from pooled tissue from testis, ovary,
                     endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT       49 a     64 c     45 g     77 t
ORIGIN

Alignment Scores:
Pred. No.:              2.04e-24        Length:         235
Score:                  288.00          Matches:        55
Percent Similarity:     93.65%          Conservative:   4
Best Local Similarity:  87.30%          Mismatches:     4
Query Match:            52.46%          Indels:         0
DB:                     10              Gaps:           0

US-09-930-312-2 (1-107) x AW436170 (1-235)

Qy     45 PhePhePheIleValPheSerAspAlaIleCysTrpIleProValPheValValLysIle 64
          ||||||||||||   ||||||||||||||||||||||||||||||||||:::||||||
Db      8 TTTTTTTTTATAGGGTTCTCTGATGCCATCTGCTGGATTCCCGTATTTGTCATTAAAATT 67

Qy     65 LeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIleValIlePhePhe 84
          ||||||||||||||||||||||||||   ::::::||||||||   ||||||||||||
Db     68 CTCTCCCTCTTCCGGGTGGAAATACCAGGCTCACTCACTTCCTGGGGGGTCATTTTTTTC 127
```

```
Qy      85 LeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThrAsnPhePheLys 104
           ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db     128 CTTCCGGTGAACAGTGCCTTGAACCCAATCCTTTACACTCTCACAACCAGCTTTTTCAAG 187

Qy     105 AspLysLeu 107
           |||   |||
Db     188 GACACGTTG 196

RESULT 5
BF021857
LOCUS       BF021857                456 bp    mRNA    linear   EST 29-DEC-2000
DEFINITION  uy58h08.y1 McCarrey Eddy round spermatid Mus musculus cDNA clone
            IMAGE:3663807 5' similar to TR:Q9VYG0 Q9VYG0 CG4187 PROTEIN. ;,
            mRNA sequence.
ACCESSION   BF021857
VERSION     BF021857.1  GI:10753189
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 456)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:1424575
            Seq primer: Primer name ambiguous
            High quality sequence stop: 386.
FEATURES             Location/Qualifiers
     source          1. .456
                     /organism="Mus musculus"
                     /strain="CD-1"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3663807"
                     /clone_lib="McCarrey Eddy round spermatid"
                     /sex="male"
                     /tissue_type="round spermatids, pooled from multiple mice"
                     /dev_stage="60 day"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: testis; Vector: pBluescript SK+ (Stratagene
                     ); Site_1: XhoII; Site_2: EcoRI; cDNA oligo dT-primed
                     [5'-(GA)10-ACTAGTCTCGAGTTTTTTTTTTTTT-3'] and directionally
                     cloned  using 5' linkers 5'-AATTCGGCACGAG-3' and
                     5'-CTCGTGCCG-3'. Size selection of >400bp material gives
                     average insert size ranging from 1-2 kb. Library was mass
                     excised (from  lambda-UniZAP-XR) and resulting
                     single-stranded phagemids were prepped and tranformed
                     into DH10B. Library contains 98.5% recombinants.
                     References: J. Androl. 20:635-639 and Gene 25:263-269.
                     Library constructed and donated by J. McCarrey, Ph.D.
                     (Southwest Foundation for Biomedical Research, Dept. of
                     Genetics); excision done by E.M. Eddy, Ph.D. (National
                     Institutes of  Health, National Institute of Environmental
                     Health Sciences).  Original lambda-based library is
                     available through ATCC, catalog #63423."
BASE COUNT      120 a    110 c    111 g    115 t
ORIGIN

Alignment Scores:
Pred. No.:              2.46e-09        Length:         456
Score:                  162.00          Matches:        30
Percent Similarity:     100.00%         Conservative:   3
Best Local Similarity:  90.91%          Mismatches:     0
Query Match:            29.51%          Indels:         0
DB:                     12              Gaps:           0

US-09-930-312-2 (1-107) x BF021857 (1-456)

Qy      75 ThrMetThrSerTrpIleValIlePhePheLeuProValAsnSerAlaLeuAsnProIle 94
           |||:::|||||||||||||||||:::|||||||||||||||||||||||||||||||||||
Db      40 ACAATCACTTCCTGGATCGTGGTTTTTTTCCTTCCGGTGAACAGCGCCTTAAACCCCATC 99

Qy      95 LeuTyrThrLeuThrThrAsnPhePheLysAspLysLeu 107
           |||||||||||||||||:::|||||||||||||||||||
Db     100 CTCTACACTCTGACGACCTCCTTTTTTAAGGACAAGTTG 138

RESULT 6
BB632910
LOCUS       BB632910                574 bp    mRNA    linear   EST 26-OCT-2001
DEFINITION  BB632910 RIKEN full-length enriched, adult male hypothalamus Mus
            musculus cDNA clone A230094D06 5', mRNA sequence.
ACCESSION   BB632910
VERSION     BB632910.1  GI:16469432
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 574)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
            , Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
            ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
```

```
     source          1. .574
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="A230094D06"
                     /clone_lib="RIKEN full-length enriched, adult male
                     hypothalamus"
                     /sex="male"
                     /tissue_type="hypothalamus"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 20.0 and subtraction to Rot = 459.0. Second
                     strand cDNA was prepared with the primer adapter of
                     sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC
                     3']. cDNA was cleaved with XhoI and BamHI. Vector: a
                     modified pBluescript KS(+) after bulk excision from Lambda
                     FLC I. "
BASE COUNT      153 a    139 c    141 g    141 t
ORIGIN

Alignment Scores:
Pred. No.:                3.09e-09       Length:          574
Score:                    162.00         Matches:         30
Percent Similarity:       100.00%        Conservative:    3
Best Local Similarity:    90.91%         Mismatches:      0
Query Match:              29.51%         Indels:          0
DB:                       10             Gaps:            0

US-09-930-312-2 (1-107) x BB632910 (1-574)

Qy      75 ThrMetThrSerTrpIleValIlePhePheLeuProValAsnSerAlaLeuAsnProIle 94
           |||:::|||||||||||||||:::|||||||||||||||||||||||||||||||||||
Db      70 ACAATCACTTCCTGGATCGTGGTTTTTTTCCTTCCGGTGAACAGCGCCTTAAACCCCATC 129

Qy      95 LeuTyrThrLeuThrThrAsnPhePheLysAspLysLeu 107
           |||||||||||||||||:::|||||||||||||||||||
Db     130 CTCTACACTCTGACGACCTCCTTTTTTAAGGACAAGTTG 168

RESULT 7
BI630907
LOCUS       BI630907                573 bp    mRNA    linear   EST 10-SEP-2001
DEFINITION  RH60250.5prime RH Drosophila melanogaster normalized Head pFlc-1
            Drosophila melanogaster cDNA clone RH60250 5 similar to rk:
            FBan0008930 GO:[plasma membrane (GO:0005886); G protein linked
            receptor (GO:0004930); G protein linked receptor (GO:0004930)]
            located on: 2L 34E2-34E2;: 08/23/2001, mRNA sequence.
ACCESSION   BI630907
VERSION     BI630907.1  GI:15533117
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 573)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RH Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003642: arm:2L [13845947,14118016]
            estimated-cyto:34E1-34F2: 08/23/2001
            Plate: RH.602  row: E  column: 2
            High quality sequence stop: 540.
FEATURES              Location/Qualifiers
     source           1. .573
                      /organism="Drosophila melanogaster"
                      /db_xref="taxon:7227"
                      /clone="RH60250"
                      /clone_lib="RH Drosophila melanogaster normalized Head
                      pFlc-1"
                      /sex="male and female"
                      /dev_stage="Adult"
                      /lab_host="DH5-alpha TonA"
                      /note="Organ: head; Vector: pFlc1; Site_1: XhoI; Site_2:
                      BamHI; Library was kindly generated by Piero Carninci at
                      the RIKEN. The library  was normalized and excised using
                      Cre recombinase. Plasmid cDNA library."
BASE COUNT      129 a    148 c    157 g    139 t
ORIGIN

Alignment Scores:
Pred. No.:                6.89e-07       Length:          573
Score:                    142.00         Matches:         33
Percent Similarity:       52.38%         Conservative:    22
Best Local Similarity:    31.43%         Mismatches:      40
Query Match:              25.87%         Indels:          10
DB:                       13             Gaps:            3

US-09-930-312-2 (1-107) x BI630907 (1-573)

Qy       2 ValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSerIle 21
             :::|||      |||||||||   ::::::      |||:::   |||:::   :::|||
Db      26 ATCAACGGATGCGCATTTCTCACTCTGATGGGTTGCTACCTGAAGATGTACTGGGCCATA 85

Qy      22 Gln---LysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
             :::      |||      |||   :::   |||
Db      86 AGGGGCAGCCAGGCATGGAACACGAATGATTCGCGA---------------------- 121

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
             :::|||   |||          :::|||||:::|||   :::|||||   |||:::
Db     122 ATTGCCAAGCGAATGGCTCTGTTGGTCTTTACTGACTTCCTCTGCTGGTCACCTATCGCC 181

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
                |||   :::::::|||   :::::::::                :::   |||
Db     182 TTCTTCTCGATCACTGCCATCTTTGGCCTGCAACTGATTTCGCTGGAGCAGGCCAAGATC 241

Qy      81 ---ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThr 99
               :::|||   |||||||:::||||||      ||||||   ||||||   :::
Db     242 TTTACGGTATTTGTGCTGCCCTTGAATAGTTGCTGCAATCCCTTTCTCTACGCCATAATG 301

Qy     100 ThrAsnPhePheLys 104
           |||      ||||||
Db     302 ACCAAGCAGTTCAAG 316

RESULT 8
BI628622
LOCUS       BI628622                610 bp    mRNA    linear   EST 10-SEP-2001
DEFINITION  RH57122.5prime RH Drosophila melanogaster normalized Head pFlc-1
            Drosophila melanogaster cDNA clone RH57122 5 similar to rk:
            FBan0008930 GO:[plasma membrane (GO:0005886); G protein linked
            receptor (GO:0004930); G protein linked receptor (GO:0004930)]
            located on: 2L 34E2-34E2;: 08/22/2001, mRNA sequence.
ACCESSION   BI628622
VERSION     BI628622.1  GI:15530832
KEYWORDS    EST.
SOURCE      fruit fly.
```

```
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 610)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RH Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003642: arm:2L [13845947,14118016]
            estimated-cyto:34E1-34F2: 08/22/2001
            Plate: RH.571  row: B  column: 10
            High quality sequence stop: 539.
FEATURES             Location/Qualifiers
     source          1. .610
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RH57122"
                     /clone_lib="RH Drosophila melanogaster normalized Head
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="Adult"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: head; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library  was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT      143 a    160 c    166 g    139 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:                7.33e-07        Length:         610
Score:                    142.00          Matches:        33
Percent Similarity:       52.38%          Conservative:   22
Best Local Similarity:    31.43%          Mismatches:     40
Query Match:              25.87%          Indels:         10
DB:                       13              Gaps:           3

US-09-930-312-2 (1-107) x BI628622 (1-610)

Qy       2 ValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSerIle 21
           :::|||      |||||||||   ::::::      |||:::   |||:::   :::|||
Db      26 ATCAACGGATGCGCATTTCTCACTCTGATGGGTTGCTACCTGAAGATGTACTGGGCCATA 85

Qy      22 Gln---LysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAla 40
           :::       |||       |||   :::   |||
Db      86 AGGGGCAGCCAGGCATGGAACACGAATGATTCGCGA---------------------- 121

Qy      41 ValAlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPhe 60
           :::|||   |||         :::||||||:::|||   :::||||||   |||:::
Db     122 ATTGCCAAGCGAATGGCTCTGTTGGTCTTTACTGACTTCCTCTGCTGGTCACCTATCGCC 181

Qy      61 ValValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIle 80
                 |||   :::::|||   :::::::::               :::   |||
Db     182 TTCTTCTCGATCACTGCCATCTTTGGCCTGCAACTGATTTCGCTGGAGCAGGCCAAGATC 241

Qy      81 ---ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThr 99
              :::|||   ||||||:::||||||      ||||||   ||||||   :::
Db     242 TTTACGGTATTTGTGCTGCCCTTGAATAGTTGCTGCAATCCCTTTCTCTACGCCATAATG 301

Qy     100 ThrAsnPhePheLys 104
           |||      ||||||
Db     302 ACCAAGCAGTTCAAG 316
```

```
RESULT 9
BM960298
LOCUS       BM960298                511 bp    mRNA    linear   EST 18-MAR-2002
DEFINITION  cihA4L14S Ascidian hemocytes cDNA library Ciona intestinalis cDNA,
            mRNA sequence.
ACCESSION   BM960298
VERSION     BM960298.1  GI:19543725
KEYWORDS    EST.
SOURCE      Ciona intestinalis.
  ORGANISM  Ciona intestinalis
            Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
            Phlebobranchia; Cionidae; Ciona.
REFERENCE   1  (bases 1 to 511)
  AUTHORS   Shida,K., Terajima,D., Satake,M., Kawazoe,Y. and Kasuya,A.
  TITLE     Sequence tags expressed in hemocytes of Ciona intestinalis
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Masanobu Satake
            Department of Molecular Immunology
            Institute of Development, Aging and Cancer, Tohoku University
            Seiryo-machi 4-1, Aoba-ku, Sendai 980-8575, Japan
            Tel: 81-22-717-8477
            Fax: 81-22-717-8482
            Email: satake@idac.tohoku.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .511
                     /organism="Ciona intestinalis"
                     /db_xref="taxon:7719"
                     /clone_lib="Ascidian hemocytes cDNA library"
                     /cell_type="hemocytes"
                     /dev_stage="adult"
                     /note="Vector: pZL1; Site_1: EcoRI"
BASE COUNT      119 a    118 c    124 g    150 t
ORIGIN

Alignment Scores:
Pred. No.:                2.08e-06        Length:         511
Score:                    137.50          Matches:        26
Percent Similarity:       52.08%          Conservative:   24
Best Local Similarity:    27.08%          Mismatches:     45
Query Match:              25.05%          Indels:         1
DB:                       14              Gaps:           1

US-09-930-312-2 (1-107) x BM960298 (1-511)

Qy       3 AsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSerIleGln 22
           |||      ||||||:::   :::      ||||||:::   ::::::         :::
Db      40 AATTTTACCGCGTTTGTGTTTGTTGCGTCATCATACGTCATCATCTACGTCAAAGCGCGT 99

Qy      23 LysThrAlaLeuGln---ThrThrGluValArgAsnCysPheGlyArgGluValAlaVal 41
           :::   |||:::                                  :::      |||:::
Db     100 CGATACGCGATGCACTTCGATATCAGTTCTCCCAGGGTGATGCTACAACATAGAGCCATG 159

Qy      42 AlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPheVal 61
           :::         ::::::   :::|||      ||||||:::|||:::   :::
Db     160 CAAAGGAAGATCATCAGATTAATCGTCACCGACTTTTGTTGTTGGGTTCCCATCTCCATC 219

Qy      62 ValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIleVal 81
              :::            :::|||:::||||||   |||   :::
Db     220 ACAGTGTTCGTGAGGTTTAGCGGCATCGAGGTTCCCGATATCATGTACGCGGTGTCTGCG 279

Qy      82 IlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThr 97
           :::      ||||||:::||||||||||||||||||||||||||||:::
Db     280 GTCGTCCTTCTCCCGCTGAATTCCGCGCTGAACCCGATCCTTTACTCG 327

RESULT 10
AV995043
LOCUS       AV995043                522 bp    mRNA    linear   EST 15-MAR-2002
DEFINITION  AV995043 Nori Satoh unpublished cDNA library, tailbud embryo Ciona
            intestinalis cDNA clone citb39n21 5', mRNA sequence.
ACCESSION   AV995043
VERSION     AV995043.1  GI:19486377
KEYWORDS    EST.
```

```
SOURCE      Ciona intestinalis.
  ORGANISM  Ciona intestinalis
            Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
            Phlebobranchia; Cionidae; Ciona.
REFERENCE   1  (bases 1 to 522)
  AUTHORS   Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
  TITLE     Expressed genes in Ciona intestinalis
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Nori Satoh
            Department of Zoology
            Kyoto University
            Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
            Tel: 81-75-753-4081
            Fax: 81-75-705-1113
            Email: satoh@ascidian.zool.kyoto-u.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .522
                     /organism="Ciona intestinalis"
                     /db_xref="taxon:7719"
                     /clone="citb39n21"
                     /clone_lib="Nori Satoh unpublished cDNA library, tailbud
                     embryo"
                     /tissue_type="whole animal"
                     /dev_stage="tailbud embryo"
                     /note="Vector: pBluescript SK"
BASE COUNT      121 a    122 c    127 g    152 t
ORIGIN

Alignment Scores:
Pred. No.:                2.12e-06       Length:          522
Score:                    137.50         Matches:         26
Percent Similarity:       52.08%         Conservative:    24
Best Local Similarity:    27.08%         Mismatches:      45
Query Match:              25.05%         Indels:          1
DB:                       10             Gaps:            1

US-09-930-312-2 (1-107) x AV995043 (1-522)

Qy       3 AsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSerIleGln 22
           |||      ||||||:::   :::      ||||||:::   ::::::          :::
Db      79 AATTTTACCGCGTTTGTGTTTGTTGCGTCATCATACGTCATCATCTACGTCAAAGCGCGT 138

Qy      23 LysThrAlaLeuGln---ThrThrGluValArgAsnCysPheGlyArgGluValAlaVal 41
           :::   |||:::                          :::      |||:::
Db     139 CGATACGCGATGCACTTCGATATCAGTTCTCCCAGGGTGATGCTACAACATAGAGCCATG 198

Qy      42 AlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPheVal 61
             :::         ::::::   :::|||      ||||||:::|||:::   :::
Db     199 CAAAGGAAGATCATCAGATTAATCGTCACCGACTTTTGTTGTTGGGTTCCCATCTCCATC 258

Qy      62 ValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIleVal 81
               :::            :::|||:::||||||   |||   :::
Db     259 ACAGTGTTCGTGAGGTTTAGCGGCATCGAGGTTCCCGATATCATGTACGCGGTGTCTGCG 318

Qy      82 IlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThr 97
           :::      ||||||:::||||||||||||||||||||||||||||:::
Db     319 GTCGTCCTTCTCCCGCTGAATTCCGCGCTGAACCCGATCCTTTACTCG 366

RESULT 11
AV883650/c
LOCUS       AV883650                585 bp    mRNA    linear   EST 08-NOV-2001
DEFINITION  AV883650 Nori Satoh unpublished cDNA library, tailbud embryo Ciona
            intestinalis cDNA clone rcitb39n21 3', mRNA sequence.
ACCESSION   AV883650
VERSION     AV883650.1  GI:16871174
KEYWORDS    EST.
SOURCE      Ciona intestinalis.
  ORGANISM  Ciona intestinalis
            Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
            Phlebobranchia; Cionidae; Ciona.
REFERENCE   1  (bases 1 to 585)
  AUTHORS   Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
  TITLE     Expressed genes in Ciona intestinalis
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Nori Satoh
            Department of Zoology
            Kyoto University
            Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
            Tel: 81-75-753-4081
            Fax: 81-75-705-1113
            Email: satoh@ascidian.zool.kyoto-u.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .585
                     /organism="Ciona intestinalis"
                     /db_xref="taxon:7719"
                     /clone="rcitb39n21"
                     /clone_lib="Nori Satoh unpublished cDNA library, tailbud
                     embryo"
                     /tissue_type="whole animal"
                     /dev_stage="tailbud embryo"
                     /note="Vector: pBluescript SK"
BASE COUNT      175 a    137 c    132 g    141 t
ORIGIN

Alignment Scores:
Pred. No.:                2.37e-06       Length:          585
Score:                    137.50         Matches:         26
Percent Similarity:       52.08%         Conservative:    24
Best Local Similarity:    27.08%         Mismatches:      45
Query Match:              25.05%         Indels:          1
DB:                       10             Gaps:            1

US-09-930-312-2 (1-107) x AV883650 (1-585)

Qy       3 AsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSerIleGln 22
           |||      ||||||:::   :::      ||||||:::   ::::::          :::
Db     503 AATTTTACCGCGTTTGTGTTTGTTGCGTCATCATACGTCATCATCTACGTCAAAGCGCGT 444

Qy      23 LysThrAlaLeuGln---ThrThrGluValArgAsnCysPheGlyArgGluValAlaVal 41
           :::   |||:::                          :::      |||:::
Db     443 CGATACGCGATGCACTTCGATATCAGTTCTCCCAGGGTGATGCTACAACATAGAGCCATG 384

Qy      42 AlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPheVal 61
             :::         ::::::   :::|||      ||||||:::|||:::   :::
Db     383 CAAAGGAAGATCATCAGATTAATCGTCACCGACTTTTGTTGTTGGGTTCCCATCTCCATC 324

Qy      62 ValLysIleLeuSerLeuPheArgValGluIleProAspThrMetThrSerTrpIleVal 81
               :::            :::|||:::||||||   |||   :::
Db     323 ACAGTGTTCGTGAGGTTTAGCGGCATCGAGGTTCCCGATATCATGTACGCGGTGTCTGCG 264

Qy      82 IlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThr 97
           :::      ||||||:::||||||||||||||||||||||||||||:::
Db     263 GTCGTCCTTCTCCCGCTGAATTCCGCGCTGAACCCGATCCTTTACTCG 216

RESULT 12
BB553437
LOCUS       BB553437                676 bp    mRNA    linear   EST 26-OCT-2001
DEFINITION  BB553437 RIKEN full-length enriched, 2 days pregnant adult female
            ovary Mus musculus cDNA clone E330007C19 3' similar to M81310 Mouse
            luteinizing hormone receptor mRNA, mRNA sequence.
ACCESSION   BB553437
VERSION     BB553437.2  GI:16448285
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 676)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
            , Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
```

```
  JOURNAL   Unpublished (2001)
COMMENT     On Aug 1, 2000 this sequence version replaced gi:9639803.
            Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
            ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .676
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="E330007C19"
                     /clone_lib="RIKEN full-length enriched, 2 days pregnant
                     adult female ovary"
                     /sex="female"
                     /tissue_type="ovary"
                     /dev_stage="2 days pregnant adult"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAGCGGCCGCAACTCGAGTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. Second strand cDNA was prepared with the
                     primer adapter of sequence [5'
                     GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA
                     was cleaved with BamHI and XhoI. Vector: a modified
                     pBluescript KS(+) after bulk excision from Lambda FLC I."
BASE COUNT      168 a    177 c    124 g    207 t
ORIGIN

Alignment Scores:
Pred. No.:              1.06e-05        Length:         676
Score:                  132.50          Matches:        28
Percent Similarity:     50.96%          Conservative:   25
Best Local Similarity:  26.92%          Mismatches:     44
Query Match:            24.13%          Indels:         7
DB:                     10              Gaps:           2

US-09-930-312-2 (1-107) x BB553437 (1-676)

Qy       2 ValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSerIle 21
           :::|||   :::||||||::::::|||         |||:::   ::::::   ::::::
Db     148 CTCAATGCAGTGGCCTTTGTCGTCATCTGTGCTTGCTACGTTAGGATATACTTTGCAGTT 207

Qy      22 GlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAlaVal 41
           |||         |||                              ::::::      :::
Db     208 CAAAATCCAGAGCTGACGGCTCCT----------------AACAAGGACACAAAAATT 249

Qy      42 AlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPheVal 61
           |||   :::         ::::::|||:::|||      |||      |||:::
Db     250 GCTAAGAAGATGGCCATCCTCATCTTCACAGACTTCACATGCATGGCACCCATCTCATTC 309

Qy      62 ValLysIleLeuSerLeuPheArgValGlu---IleProAspThrMetThrSerTrpIle 80
             |||   :::   |||:::|||      |||      |||   :::      :::
Db     310 TTTGCCATCTCAGCTGCCTTCAAAGTACCCCTTATCACTGTCACCAACTCAAAAGTTCTG 369

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           ::::::   |||   ||||||||||||      ||||||   ||||||   :::   |||
Db     370 CTGGTCCTTTTTTATCCTGTCAATTCTTGTGCCAACCCATTTCTGTACGCAGTGTTCACG 429

Qy     101 AsnPhePheLys 104
             |||:::
Db     430 AAGGCATTTCAG 441

RESULT 13
BM454054
LOCUS       BM454054               1113 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6419563 NCI_CGAP_Ov44 Mus musculus cDNA clone
            IMAGE:5504347 5', mRNA sequence.
ACCESSION   BM454054
VERSION     BM454054.1  GI:18503094
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1113)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Aaron Hsueh
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12145  row: i  column: 20
            High quality sequence start: 14
            High quality sequence stop: 702.
FEATURES             Location/Qualifiers
     source          1. .1113
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5504347"
                     /clone_lib="NCI_CGAP_Ov44"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: ovary, PMSG-treated; Vector:
                     pCMV-SPORT6.ccdb; Site_1: EcoRV; Site_2: NotI; Cloned
                     unidirectionally. Primer: Oligo dT. Average insert size
                     2.2 kb. Library constructed by Life Technologies. Note:
                     this is a NCI_CGAP Library."
BASE COUNT      318 a    253 c    180 g    362 t
ORIGIN
```

```
Alignment Scores:
Pred. No.:              1.72e-05        Length:         1113
Score:                  132.50          Matches:        28
Percent Similarity:     50.96%          Conservative:   25
Best Local Similarity:  26.92%          Mismatches:     44
Query Match:            24.13%          Indels:         7
DB:                     13              Gaps:           2

US-09-930-312-2 (1-107) x BM454054 (1-1113)

Qy       2 ValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSerIle 21
           :::|||   :::||||||:::::::|||        |||:::   ::::::   ::::::
Db      34 CTCAATGCAGTGGCCTTTGTTGTCATCTGTGCTTGCTACGTTAGGATATACTTTGCAGTT 93

Qy      22 GlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAlaVal 41
           |||         |||                        ::::::      :::
Db      94 CAAAATCCAGAGCTGACGGCTCCT-----------------AACAAGGACACAAAAATT 135

Qy      42 AlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPheVal 61
           |||   :::         :::::::|||:::|||      |||      |||:::
Db     136 GCTAAGAAGATGGCCATCCTCATCTTCACAGACTTCACATGCATGGCGCCCATCTCATTC 195

Qy      62 ValLysIleLeuSerLeuPheArgValGlu---IleProAspThrMetThrSerTrpIle 80
             |||   :::   |||:::|||      |||      |||   :::      :::
Db     196 TTTGCCATCTCAGCTGCCTTCAAAGTACCCCTTATCACTGTCACCAACTCAAAAGTTCTG 255

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           ::::::   |||   ||||||||||||      ||||||   ||||||   :::   |||
Db     256 CTGGTCCTTTTTTATCCTGTCAATTCTTGTGCCAACCCATTTCTGTACGCAGTGTTCACG 315

Qy     101 AsnPhePheLys 104
             |||:::
Db     316 AAGGCATTTCAG 327

RESULT 14
BM461149
LOCUS       BM461149               1142 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6419913 NCI_CGAP_Ov44 Mus musculus cDNA clone
            IMAGE:5503889 5', mRNA sequence.
ACCESSION   BM461149
VERSION     BM461149.1  GI:18510189
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1142)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Aaron Hsueh
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12144  row: f  column: 18
            High quality sequence start: 16
            High quality sequence stop: 685.
FEATURES             Location/Qualifiers
     source          1. .1142
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5503889"
                     /clone_lib="NCI_CGAP_Ov44"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: ovary, PMSG-treated; Vector:
                     pCMV-SPORT6.ccdb; Site_1: EcoRV; Site_2: NotI; Cloned
                     unidirectionally. Primer: Oligo dT. Average insert size
                     2.2 kb. Library constructed by Life Technologies. Note:
                     this is a NCI_CGAP Library."
BASE COUNT      313 a    269 c    210 g    350 t
ORIGIN

Alignment Scores:
Pred. No.:              1.77e-05        Length:         1142
Score:                  132.50          Matches:        28
Percent Similarity:     50.96%          Conservative:   25
Best Local Similarity:  26.92%          Mismatches:     44
Query Match:            24.13%          Indels:         7
DB:                     13              Gaps:           2

US-09-930-312-2 (1-107) x BM461149 (1-1142)

Qy       2 ValAsnLeuLeuAlaPheLeuIleIleValPheSerTyrIleThrMetPheCysSerIle 21
           :::|||   :::||||||:::::::|||        |||:::   ::::::   ::::::
Db     109 CTCAATGCAGTGGCCTTTGTTGTCATCTGTGCTTGCTACGTTAGGATATACTTTGCAGTT 168

Qy      22 GlnLysThrAlaLeuGlnThrThrGluValArgAsnCysPheGlyArgGluValAlaVal 41
           |||         |||                        ::::::      :::
Db     169 CAAAATCCAGAGCTGACGGCTCCT-----------------AACAAGGACACAAAAATT 210

Qy      42 AlaAsnArgPhePhePheIleValPheSerAspAlaIleCysTrpIleProValPheVal 61
           |||   :::         :::::::|||:::|||      |||      |||:::
Db     211 GCTAAGAAGATGGCCATCCTCATCTTCACAGACTTCACATGCATGGCGCCCATCTCATTC 270

Qy      62 ValLysIleLeuSerLeuPheArgValGlu---IleProAspThrMetThrSerTrpIle 80
             |||   :::   |||:::|||      |||      |||   :::      :::
Db     271 TTTGCCATCTCAGCTGCCTTCAAAGTACCCCTTATCACTGTCACCAACTCAAAAGTTCTG 330

Qy      81 ValIlePhePheLeuProValAsnSerAlaLeuAsnProIleLeuTyrThrLeuThrThr 100
           ::::::   |||   ||||||||||||      ||||||   ||||||   :::   |||
Db     331 CTGGTCCTTTTTTATCCTGTCAATTCTTGTGCCAACCCATTTCTGTACGCAGTGTTCACG 390

Qy     101 AsnPhePheLys 104
             |||:::
Db     391 AAGGCATTTCAG 402

RESULT 15
AK013534
LOCUS       AK013534               2372 bp    mRNA    linear   HTC 19-JAN-2002
DEFINITION  Mus musculus adult male hippocampus cDNA, RIKEN full-length
            enriched library, clone:2900016D13:adenosine A3 receptor, full
            insert sequence.
ACCESSION   AK013534
VERSION     AK013534.1  GI:12850937
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (strain:C57BL/6J) adult male hippocampus cDNA to mRNA,
            clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:2900016D13.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
```

Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G., Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S. and Hayashizaki,Y.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409 (6821), 685-690 (2001)
MEDLINE 21085660
PUBMED 11217851
REFERENCE 5 (bases 1 to 2372)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Gemomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

FEATURES Location/Qualifiers
source 1. .2372
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:2900016D13"
/db_xref="MGD:MGI:1905754"
/db_xref="taxon:10090"
/clone="2900016D13"
/sex="male"
/tissue_type="hippocampus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
gene 1. .2372
/gene="Adora3"
misc_feature 1. .2372
/gene="Adora3"
/note="adenosine A3 receptor
data source:MGD, source key:MGI:104847, evidence:ISS"
/db_xref="MGD:MGI:104847"
BASE COUNT 549 a 578 c 583 g 662 t
ORIGIN

Alignment Scores:

Pred. No.:	0.00014	Length:	2372
Score:	127.50	Matches:	34
Percent Similarity:	47.01%	Conservative:	21
Best Local Similarity:	29.06%	Mismatches:	41
Query Match:	23.22%	Indels:	22
DB:	11	Gaps:	3

US-09-930-312-2 (1-107) x AK013534 (1-2372)

```
Qy       9 IleIleValPheSerTyrIleThr--------------------MetPheCysSerIle 21
           ::::::   ||||||::::::|||                     :::   |||   |||
Db     909 ATGGTCTTCTTCAGCTTCGTCACCTGGATCCTCGTCCCCCTGGTTGTCATGTGTGTCATC 968

Qy      22 Gln----------------------------------------LysThrAlaLeuGlnThr 28
                                                       ||||||   |||   :::
Db     969 TACCTAGACATCTTTTTACATCATCCCAAATAAGGCTCAGTCAAAACCTGTCTGGTTTCA 1028

Qy      29 ThrGluValArgAsnCysPheGlyArgGluValAlaValAlaAsnArgPhePhePheIle 48
              :::   |||          |||||||||          |||          |||   :::
Db    1029 AA-AAGACGCGTGCTTTTAATGGACGGGAGTTCAAGACAGCTAAGTCCCTGTTTATGGTT 1087

Qy      49 ValPheSerAspAlaIleCysTrpIleProValPheValValLysIleLeuSerLeuPhe 68
           :::|||      |||:::||||||:::|||:::   ::::::      :::|||   |||
Db    1088 CTCTTCTTGTTTGCGCTGTGCTGGCTGCCTTTGTCCATCATCAATTTTGTTTCCTATTTT 1147

Qy      69 ArgValGluIleProAspThrMetThrSerTrpIleValIlePhePheLeuProValAsn 88
              |||:::|||||||||   :::         :::   |||                |||
Db    1148 GATGTAAAGATACCAGAT---GTCGCAATGTGCCTGGGGATCCTGTTGTCCCACGCGAAC 1204

Qy      89 SerAlaLeuAsnProIleLeuTyrThrLeuThrThrAsnPhePheLysAsp 105
           |||   :::|||||||||:::|||                ||||||:::
Db    1205 TCCATGATGAACCCTATTGTCTACGCCTGCAAAATAAAAAAGTTCAAAGAG 1255
```

Search completed: May 26, 2003, 21:20:09
Job time : 1725 secs